

## **Protein Sequence Searches - February 2005**

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

**When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2005, 21:44:23 ; Search time 136 Seconds  
(without alignments)  
951.880 Million cell updates/sec

Title: US-10-031-918A-1

Perfect score: 1997

Sequence: 1 MVSVSIRKQAEGPATIL.....LFGFGPLTETVLSVAI 389

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1424015 seqs, 332791073 residues

Total number of hits satisfying chosen parameters: 1424015

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep:\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep:\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1994	99.8	389	16	US-10-450-230-1
2	1994	99.8	389	16	US-10-450-230-10
3	1969	98.6	389	16	US-10-659-755-7
4	1933	96.8	389	17	US-10-491-823-316
5	1911	95.7	389	16	US-10-659-755-9
6	1903	95.3	389	17	US-10-491-823-314
7	1840	92.1	389	17	US-10-491-823-320
8	1839	92.1	405	15	US-10-425-114-52518
9	1820	91.1	386	15	US-10-424-599-148762
10	1809.5	90.6	388	15	US-10-424-599-171157
11	1808.5	90.6	388	15	US-10-424-599-171068
12	1804.5	90.4	388	15	US-10-424-599-171091
13	1803.5	90.3	388	15	US-10-424-599-171202

14	1798.5	90.1	388	15	US-10-424-599-171113	Sequence 171113,
15	1795.5	89.9	388	15	US-10-424-599-171124	Sequence 171124,
16	1761	88.2	389	16	US-10-641-495-5	Sequence 5, Appli
17	1734	86.8	389	14	US-10-136-444-2	Sequence 2, Appli
18	1733	86.8	391	15	US-10-424-599-169160	Sequence 169160,
19	1719	86.1	387	15	US-10-425-114-50078	Sequence 50078, A
20	1715.5	85.9	384	15	US-10-424-599-148763	Sequence 148763,
21	1715.5	85.4	396	16	US-10-450-230-9	Sequence 9, Appli
22	1696	84.9	392	15	US-10-224-493-2	Sequence 2, Appli
23	1692	84.7	392	15	US-10-224-493-6	Sequence 6, Appli
24	1690	84.6	392	15	US-10-224-493-4	Sequence 4, Appli
25	1686.5	84.5	397	15	US-10-425-114-40814	Sequence 40814, A
26	1676.5	84.0	401	15	US-10-425-114-37280	Sequence 37280, A
27	1672.5	83.8	390	15	US-10-424-599-171190	Sequence 171190,
28	1671	83.7	391	17	US-10-491-823-324	Sequence 324, App
29	1660.5	83.1	395	16	US-10-715-872-168	Sequence 168, App
30	1634	81.8	398	16	US-10-437-963-165957	Sequence 165957,
31	1630	81.6	394	14	US-10-289-757-119	Sequence 119, App
32	1630	81.6	398	14	US-10-289-757-186	Sequence 186, App
33	1630	81.6	398	15	US-10-431-273-76	Sequence 76, Appl
34	1627	81.5	398	15	US-10-431-273-77	Sequence 77, Appl
35	1550	77.6	309	17	US-10-491-823-24	Sequence 24, Appl
36	1544	77.3	360	9	US-09-829-378-3	Sequence 3, Appli
37	1543	77.3	382	16	US-10-437-963-165955	Sequence 165955,
38	1524	76.3	338	15	US-10-425-114-43099	Sequence 43099, A
39	1508.5	75.5	394	9	US-09-837-654-1	Sequence 1, Appli
40	1508.5	75.5	394	14	US-10-390-658-1	Sequence 1, Appli
41	1492	74.7	398	16	US-10-767-701-45589	Sequence 45589, A
42	1439.5	72.1	403	16	US-10-437-963-148991	Sequence 148991,
43	1409.5	70.6	393	16	US-10-450-230-8	Sequence 8, Appli
44	1293	64.7	360	9	US-09-829-378-4	Sequence 4, Appli
45	1243	62.2	404	16	US-10-767-701-46457	Sequence 46457, A

#### ALIGNMENTS

#### RESULT 1

US-10-450-230-1  
; Sequence 1, Application US/10450230  
; Publication No. US20040106175A1  
; GENERAL INFORMATION:  
; APPLICANT: NOEL, JOSEPH P.  
; APPLICANT: AUSTIN, MICHAEL B.  
; APPLICANT: BOWMAN, MARIANNE E.  
; TITLE OF INVENTION: METHODS OF PRODUCING POLYKETIDE SYNTHASE MUTANTS AND  
; TITLE OF INVENTION: COMPOSITIONS AND USES THEREOF  
; FILE REFERENCE: SALR2370-3  
; CURRENT APPLICATION NUMBER: US/10/450,230  
; CURRENT FILING DATE: 2003-11-26  
; PRIOR FILING DATE: 2001-12-14  
; PRIOR FILING DATE: 2003-11-26  
; PRIOR FILING DATE: 2000-12-15  
; NUMBER OF SEQ ID NOS: 64  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 389  
; TYPE: PRT  
; ORGANISM: Medicago sativa  
US-10-450-230-1

Query Match 99.8%; Score 1994; DB 16; Length 389;  
Best Local Similarity 99.7%; Pred. No. 7.2e-192;  
Matches 388; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MVSVSIRKQAEGPATILAI GTANPANCVCQSTYDFYFKITNSEHKTELKEKQFQMC 60  
Db 1 MVSVSIRKQAEGPATILAI GTANPANCVCQSTYDFYFKITNSEHKTELKEKQFQMC 60  
QY 61 DKSMIKRRYMYLITEILKENPNVCYMAPSLDARQDMVVVPRIGKEAAVKAKEWGQP 120  
Db 61 DKSMIKRRYMYLITEILKENPNVCYMAPSLDARQDMVVVPRIGKEAAVKAKEWGQP 120

Qy	121	KSKI	THLI	IVCTT	SGVDM	PGADY	QLTKL	GLSPY	VKRY	YMY	QQG	FAGGT	VLR	AKOLA	EN	180	
Db	121	KSKI	THLI	IVCTT	SGVDM	PGADY	QLTKL	GLSPY	VKRY	YMY	QQG	FAGGT	VLR	AKOLA	EN	180	
Qy	181	NKGAR	VLVV	CVSE	VTAT	FRGSD	THLDS	LVGQA	ALF	GDG	AAAL	IVGSD	PVP	PEI	EKPI	FEMV	240
Db	181	NKGAR	VLVV	CVSE	VTAT	FRGSD	THLDS	LVGQA	ALF	GDG	AAAL	IVGSD	PVP	PEI	EKPI	FEMV	240
Qy	241	WTAQ	TIAP	DSEGA	IDGHL	REAGL	TEPHL	LKDV	PGIV	SKNI	ITKAL	VEAPE	PLGI	DYNS	IFW	300	
Db	241	WTAQ	TIAP	DSEGA	IDGHL	REAGL	TEPHL	LKDV	PGIV	SKNI	ITKAL	VEAPE	PLGI	DYNS	IFW	300	
Qy	301	IAHGG	PATLD	VEQ	EKL	ALKE	PKKNAT	REVUSE	YGN	SSAC	VFLID	EMR	KKSTQ	NGLT	360		
Db	301	IAHGG	PATLD	VEQ	EKL	ALKE	PKKNAT	REVUSE	YGN	SSAC	VFLID	EMR	KKSTQ	NGLT	360		
Qy	361	TGEGL	EWGLV	FGF	GG	FLGT	ETI	TVL	RSVAI	389							
Db	361	TGEGL	EWGLV	FGF	GG	FLGT	ETI	TVL	RSVAI	389							

RESULT 2  
US-10-450-230-10  
; Sequence 10, Application US/10450230  
; Publication No. US20040106175A1  
; GENERAL INFORMATION:  
; APPLICANT: NOEL, JOSEPH P.  
; APPLICANT: AUSTIN, MICHAEL B.  
; APPLICANT: BOWMAN, MARRIANNE E.  
; TITLE OF INVENTION: METHODS OF PRODUCING POLYKETIDE SYNTHASE MUTANTS AND  
; TITLE OF INVENTION: COMPOSITIONS AND USES THEREOF  
; FILE REFERENCE: SALK2370-3  
; CURRENT APPLICATION NUMBER: US/10/450,230  
; CURRENT FILING DATE: 2003-11-26  
; PRIOR APPLICATION NUMBER: PCT/US01/48523  
; PRIOR FILING DATE: 2001-12-14  
; PRIOR APPLICATION NUMBER: 60/255,811  
; PRIOR FILING DATE: 2000-12-15  
; NUMBER OF SEQ ID NOS: 64  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 389  
; TYPE: PRT  
; ORGANISM: Medicago sativa  
US-10-450-230-10

Db	301	IAHPGGPAILDQVEQKALALPEKKNATREVLSEYGNMSSACVLFILDEMRRKKSQTONGLKT	360
Qy	361	TGEGLEWGLVFGFGPGLTIETVILRSVAI	389
Db	361	TGEGLEWGLVFGFGPGLTIETVILRSVAI	389
RESULT 3			
US-10-659-755-7			
; Sequence 7, Application US/10659755			
; Publication No. US20040128711A1			
; GENERAL INFORMATION:			
; APPLICANT: DIXON, RICHARD A.			
; APPLICANT: LIU, CHANG-JUN			
; APPLICANT: DEAVOURS, BETTINA			
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR PRODUCTION OF FLAVONOID			
; TITLE OF INVENTION: AND ISOFLAVONOID NUTRACEUTICALS			
; FILE REFERENCE: NBLE:0070US			
; CURRENT APPLICATION NUMBER: US/10/659,755			
; CURRENT FILING DATE: 2003-09-10			
; PRIOR APPLICATION NUMBER: 60/409,447			
; PRIOR FILING DATE: 2002-09-10			
; NUMBER OF SEQ ID NOS: 14			
; SOFTWARE: Patentin Ver. 2.1			
; SEQ ID NO 7			
; LENGTH: 389			
; TYPE: PRT			
; ORGANISM: Medicago sativa			
US-10-659-755-7			
Query Match 98.6%; Score 1969; DB 16; Length 389;			
Best Local Similarity 98.5%; Pred. No. 2.4e-189;			
Matches 383; Conservative 3; Mismatches 3; Indels 0; Gaps 0;			
Qy	1	MVSVSEIRKQARAGPATILAIAGTANPANCVEQSTYDDFPFKITNSEHKTELKKEKFORMC	60
Db	1	MVSVSEIRKQARAGPATILAIAGTANPANCVEQSTYDDFPFKITNSEHKTELKKEKFORMC	60
Qy	61	DKSMIKRMYLTHEIILKENPNVCEYNAPSILDAQDMVVVPRLGKEAAVKAIKEWGQP	120
Db	61	DKSMIKRMYLTHEIILKENPNVCEYNAPSILDAQDMVVVPRLGKEAAVKAIKEWGQP	120
Qy	121	KSKITHLIVCTTSGVDMPGADYQLTLLGLRPYVKRYMMYQQGFAGGTVLRLAKDLAEN	180
Db	121	KSKITHLIVCTTSGVDMPGADYQLTLLGLRPYVKRYMMYQQGFAGGTVLRLAKDLAEN	180
Qy	181	NKGARVLVWCSEVTAVTFRPGSDTHLDSLVGQALIFDGAALIVGSDVPVPIEKPIPMV	240
Db	181	NKGARVLVWCSEVTAVTFRPGSDTHLDSLVGQALIFDGAALIVGSDVPVPIEKPIPMV	240
Qy	241	WTAQTIAPDSEGATDGHRLREAGLTFHLLKQVPGIVSKNITKALVEAPEPLGISDYNSTFW	300
Db	241	WTAQTIAPDSEGATDGHRLREAGLTFHLLKQVPGIVSKNITKALVEAPEPLGISDYNSTFW	300
Qy	301	IAHPGGPAILDQVEQKALALPEKKNATREVLSEYGNMSSACVLFILDEMRRKKSQTONGLKT	360
Db	301	IAHPGGPAILDQVEQKALALPEKKNATREVLSEYGNMSSACVLFILDEMRRKKSQTONGLKT	360
Qy	361	TGEGLEWGLVFGFGPGLTIETVILRSVAI	389
Db	361	TGEGLEWGLVFGFGPGLTIETVILRSVAI	389

RESULT 4  
US-10-491-823-316  
; Sequence 316, Application US/10491823  
; Publication No. US20050069894A1  
; GENERAL INFORMATION:  
; APPLICANT: Agriculture Victoria Services Pty Ltd  
; APPLICANT: Agresearch Limited  
; APPLICANT: SPANGENBERG, German  
; APPLICANT: SAWBRIDGE, Timothy Ivor  
; APPLICANT: ONG, Eng Kok

```

; APPLICANT: EMMERLING, Michael
; TITLE OF INVENTION: Manipulation of flavonoid biosynthesis in plants
; FILE REFERENCE: 4491-4003
; CURRENT APPLICATION NUMBER: US/10/491,823
; CURRENT FILING DATE: 2004-04-05
; PRIOR APPLICATION NUMBER: PR8113
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: PCT/AU2002/001345
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 477
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 316
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Trifolium repens
; US-10-491-823-316

Query Match          96.8%; Score 1933; DB 17; Length 389;
Best Local Similarity 96.7%; Pred. No. 1e-185;
Matches 376; Conservative 5; Mismatches 8; Indels 0; Gaps 0

Qy 1 MVSVEIRKAQRAEGPATILAIQTANPANCVQEQSTYPDFYFKITNSEHKTELKEKFORMC 60
Db 1 MVSVAEIRKAQRAEGPATILAIQTANPPNRVEQSTYPDFYFKITNSEHKTELKEKFORMC 60

Qy 61 DKSMIKERYMYLITEILLKENPNVCEYMAPSLDARQDMVVVPRLGKEAAVKAKEWGQP 120
Db 61 DKSMIKERYMYLITEILLKENPSLCEYMAPSLDARQDMVVVPRLGKEAAVKAKEWGQP 120

Qy 121 KSKITHLIVCTTSGVDMPGADYQTLKLGLRPYVKRYMYQOGXFPAGGTVLR LAKDLAEN 180
Db 121 KSKITHLIFCTTSGVDMPGADYQTLKLGLRPYVKRYMYQOGXFPAGGTVLR LAKDLAEN 180

Qy 181 NKGARVLVCESEVTAVTFRGPSTDLHDSLVGOALFGDGAALIVGSDPVPRIEPIFEMV 240
Db 181 NKGARVLVCESEVTAVTFRGPSTDLHDSLVGOALFGDGAALIVGSDPVPRIEPIFEMV 240

Qy 241 WTAQTIAPDSEGATDGHRLREAGLTFHLLKDVPGIVSKNITKALVEAPEPLGISDYSNIFW 300
Db 241 WTAQTIAPDSEGATDGHRLREAGLTFHLLKDVPGIVSKNINKALVEAFQPLGISDYSNIFW 300

Qy 301 IAHPGGPAILLDQVEQKALKPEKKNATREVLSEYGNMSSACVLFILDEMRRKKSQTONGLKT 360
Db 301 IAHPGGPAILLDQVEQKALKPEKKNATREVLSEYGNMSSACVLFILDEMRRKKSQTONGLKT 360

Qy 361 TGEGLGWGLVFGPGGGLTETVLRVSAI 389
Db 361 TGEGLDGWGLVFGPGGGLTETVLRVSAI 389

RESULT 5
US-10-659-755-9
; Sequence 9, Application US/10659755
; Publication No. US20040128711a1
; GENERAL INFORMATION:
; APPLICANT: DIXON, RICHARD A.
; APPLICANT: LIU, CHANG-JUN
; APPLICANT: DEAVOURS, BETTINA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR PRODUCTION OF FLAVONOID
; FILE REFERENCE: NBLE:007US
; CURRENT APPLICATION NUMBER: US/10/659,755
; CURRENT FILING DATE: 2003-09-10
; PRIOR FILING DATE: 2002-09-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Medicago sativa
; US-10-659-755-9

```

Db 121 KSKITHLIFCTTSGVDMFGADYQLTLLGLRPVVKRYMMYQCGCFAGGTVLRKDLAEN 180  
Qy 181 NKGARVLVVCSEVTANTFRGSDTHLDSLVGQALFGDGAALIVGSDPVEIEKPIFEW 240  
Db 181 NKGARVLVVCSEVTANTFRGSDTHLDSLVGQALFGDGAALIVGSDPVEIEKPIFEW 240  
Qy 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFELPGISDYSIFW 300  
Db 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFELPGISDYSIFW 300  
Qy 301 IAHGGPAILDQVEQKALPEKKNATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
Db 301 IAHGGPAILDQVEQKALPEKKNATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
Qy 361 TGEGLWGLVFGFGPGLTIETVLRSAI 389  
Db 361 TGEGLWGLVFGFGPGLTIETVLRSAI 389

## RESULT 7

US-10-491-823-320  
; Sequence 320, Application US/10491823  
; Publication No. US20050069884A1  
; GENERAL INFORMATION:  
; APPLICANT: Agriculture Victoria Services Pty Ltd  
; APPLICANT: AgResearch Limited  
; APPLICANT: SPANGENBERG, German  
; APPLICANT: SAWBRIDGE, Timothy Ivor  
; APPLICANT: ONG, Eng Kok  
; APPLICANT: EMMERLING, Michael  
; TITLE OF INVENTION: Manipulation of flavonoid biosynthesis in plants  
; FILE REFERENCE: 4491-4003  
; CURRENT APPLICATION NUMBER: US/10/491,823  
; CURRENT FILING DATE: 2004-04-05  
; PRIOR FILING DATE: 2004-04-05  
; PRIOR FILING DATE: 2001-10-05  
; PRIOR FILING DATE: 2002-10-04  
; NUMBER OF SEQ ID NOS: 477  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 320  
; TYPE: PRT  
; ORGANISM: Trifolium repens  
US-10-491-823-320

Query Match 92.1%; Score 1840; DB 17; Length 389;  
Best Local Similarity 89.2%; Pred. No. 2.4e-176;  
Matches 346; Conservative 28; Mismatches 14; Indels 0; Gaps 0;  
Qy 1 MVSVEIRKAQRAEGPATILAITANPANCVEQSTYDPDFYFKITNSEHKTTELKEKFORMC 60  
Db 1 MWNVEIRQAQRAEGPATILAITANPNCVQSTYDPDFYFRITNSEHKTTELKEKFORMC 60  
Qy 61 DKSIMKRYMYLTHEILKENPNVCEYMAPSLDARDQMVVVEVPRLGKEAAVKAKEWGQP 120  
Db 61 DKSIMKRYMYLTHEILKENPNVCEYMAPSLDARDQMVVVEVPRLGKEAAVKAKEWGQP 120  
Qy 121 KSKITHLIVCTTSGVDMFGADYQLTLLGLRPVVKRYMMYQCGCFAGGTVLRKDLAEN 180  
Db 121 KSKITHLIFCTTSGVDMFGADYQLTLLGLRPVVKRYMMYQCGCFAGGTVLRKDLAEN 180  
Qy 181 NKGARVLVVCSEVTANTFRGSDTHLDSLVGQALFGDGAALIVGSDPVEIEKPIFEW 240  
Db 181 NKGARVLVVCSEVTANTFRGSDTHLDSLVGQALFGDGAALIVGSDPVEIEKPIFEW 240  
Qy 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFELPGISDYSIFW 300  
Db 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFELPGISDYSIFW 300  
Qy 301 IAHGGPAILDQVEQKALPEKKNATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
Db 301 IAHGGPAILDQVEQKALPEKKNATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360

Db 301 IAHGGPAILDQVEAKLSLKPEKMQATRHVLSEYGNMSSACVLFILDEMRRKSKEDGLAT 360  
Qy 361 TGEGLWGLVFGFGPGLTIETVLRSAI 389  
Db 361 TGEGLWGLVFGFGPGLTIETVLRSAI 389  
RESULT 8  
US-10-425-114-52518  
; Sequence 52518, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 52518  
; LENGTH: 405  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3049-050-G10\_FLI.pap  
US-10-425-114-52518

Query Match 92.1%; Score 1839; DB 15; Length 405;  
Best Local Similarity 90.2%; Pred. No. 3.2e-176;  
Matches 351; Conservative 19; Mismatches 19; Indels 0; Gaps 0;  
Qy 1 MVSVEIRKAQRAEGPATILAITANPANCVEQSTYDPDFYFKITNSEHKTTELKEKFORMC 60  
Db 17 MVSVAEIRQAQRAEGPATILAITANPNRVDQSTYDPYFRITNSEHKTTELKEKFORMC 76  
Qy 61 DKSIMKRYMYLTHEILKENPNVCEYMAPSLDARDQMVVVEVPRLGKEAAVKAKEWGQP 120  
Db 77 DKSIMKRYMYLTHEILKENPNVCEYMAPSLDARDQMVVVEVPRLGKEAAVKAKEWGQP 136  
Qy 121 KSKITHLIVCTTSGVDMFGADYQLTLLGLRPVVKRYMMYQCGCFAGGTVLRKDLAEN 180  
Db 137 KSKITHLIFCTTSGVDMFGADYQLTLLGLRPVVKRYMMYQCGCFAGGTVLRKDLAEN 196  
Qy 181 NKGARVLVVCSEVTANTFRGSDTHLDSLVGQALFGDGAALIVGSDPVEIEKPIFEW 240  
Db 197 NKGARVLVVCSEVTANTFRGSDTHLDSLVGQALFGDGAALIVGSDPVEIEKPIFEW 256  
Qy 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFELPGISDYSIFW 300  
Db 257 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFELPGISDYSIFW 316  
Qy 301 IAHGGPAILDQVEQKALPEKKNATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
Db 317 IAHGGPAILDQVEQKALPEKKNATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 376  
Qy 361 TGEGLWGLVFGFGPGLTIETVLRSAI 389  
Db 377 TGEGLWGLVFGFGPGLTIETVLRSAI 405

## RESULT 9

US-10-424-599-148762  
; Sequence 148762, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua

Matches	342;	Conservative	26;	Mismatches	20;	Indels	1;	Gaps	1;
Qy	1	MVSVEIRKQRAEGPATIIAIGTANPANCVEQSTYDPDYFKYINSEHTELTKEKQFMC	60						
Db	1	MVSVEIRKQRAEGPATVMAIGTATPNCVDQSTYDPDYFRITNSEHMTLKEKFKMC	60						
Qy	61	DKSMIKRMYLVTETIILKENPNVCYMAPSLDARDQDMVVEVPRLGKAAVKAKEWGQP	120						
Db	61	DKSMIKRMYLVTETIILKENPNVCYMAPSLDARDQDMVVEVPRLGKAAVKAKEWGQP	120						
Qy	121	KSKITHLIVCTTSGVDMPGADYQLTLLGLRPYVKRYMYYQQGXPAGGTVTLRLAKDLAEN	180						
Db	121	KSKITHLIFCTTSGVDMPGADYQLTLLGLRPSVKRYMYYQQGFPAGGTVTLRLAKDLAEN	180						

QY                   NGKARVLLVVCSEVAVIFRFGPSVHLHLSLVGQALFDGAAATLVGSDPVPSEIETAFEMD 239

181                   NGKARVLLVVCSEVAVIFRFGPSVHLHLSLVGQALFDGAAATLVGSDPVPSEIETAFEMD 239

Db                   NGKARVLLVVCSEITAVIFRFGPTDHLDSL VGQALFDGAAAAVVGSDPLP-VEKPLFLQL 239

```

241 WTAQIAPFDSGALDGRKRAGLIFHLTKOVFGAVSKNIINATITKATSPFLGSDINSLW 300
240 WTAQTILPDSEGAIDGHLREVGLTFFHLKDVPGHISKNIEKALVEAFQPLGSDYNSIFW 299
301 IAHPPGPAILDQVEOKLALPEKKNATREVLSEYGNMSSACVLFILDMRKKSTQNGLKT 360
300 IAHPPGPAILDQVEAKLGLPERKEATRHVLSEYGNMSSACVLFILDMRKKSIENGLGT 359
361 TGEGLGWGLVFGFGPGLTIETVLRLSVAI 389
360 TGEGLDWGLVFGFGPGLTVEIVLRSVTV 388

RESULT 11
US-10-424-599-171068
; Sequence 171068, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684

```

```

; LENGTH: 363
; TYPE: PRNT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_12548C.1.pep
US-10-424-599-171068

Query Match          90.6%; Score 1808.5; DB 15; Length 593;
Best Local Similarity 87.9%; Pred. No. 6.8e-173;
Matches 342; Conservative 26; Mismatches 20; Indels 1; Gaps 1

Qy      1  MVSVSEIRKQARAGPATILAIAGTANPANCVEQSTYDDYFKITNSEHKTELKEKFORMC 60
Db      196 MVSVEEIRKQARAGPATVWAIGATTPNCVDQSTYDDYFRIINSEHWTELKEKFKMC 255

Qy      61  DKSMIKRRYMYLTBEILKENPNVCYMAPSLDARQDMVVVEVPRIGKEAAVKAIKEWQOP 120
Db      256 DKSMIKKRYMYLNEEILKENPSVCAYMAPSLDARQDMVVVEVPLGKEAATKAIKEWQOP 315

Qy      121 KSKITHLIVCTTSVDMPGADYQLTLLGLAPYKRYNMYQQGCFAGGTVLRKADLAEN 180
Db      316 KSKITHLIFCTSGDMPGADYQLTLLGLRPSVKRYNMYQQGCFAGGTVLRKADLAEN 375

```

QY 101 NAGARLVVVCSEIVAVIATAGFSDINLSDSVGQAFGLGAAATII VGSDFVFLIENFIFEN 219

Db 376 NKGARLVVVCSETAVTTRGPTDTHLSDVGLQALFGDGAANAIVGSDPLP VEXPLFLQLV 434





OTHER INFORMATION: Clone ID: PAT\_MRT3847\_12552C.1.pap  
US-10-424-599-171113

Query Match 90.1%; Score 1798.5; DB 15; Length 388;  
Best Local Similarity 87.4%; Pred. No. 3.7e-172; Indels 1; Gaps 1;  
Matches 340; Conservative 26; Mismatches 22; Indels 1; Gaps 1;  
QY 1 MVSVEIRKAQRAEGPATILAIATPANPCVQSTYDPDFYFKITNSEHKTTELKEKFORMC 60  
DB 1 MVSVEIRKAQRAEGPATILAIATPANPCVQSTYDPDFYFKITNSEHKTTELKEKFORMC 60  
QY 61 DKSMIKRYYNLTETILKENPNVCEYMAPSLDARDQMVVVEPRLGKEAAVKAIKEWGQP 120  
DB 61 DKSMIKRYYNLTETILKENPNVCEYMAPSLDARDQMVVVEPRLGKEAAVKAIKEWGQP 120  
QY 121 KSKITHLIVCTTSGVDMPCGADYQTLKLLGLRPSVKRYMYQQGCFAGGTVLRKDLAEN 180  
DB 121 KSKITHLIVCTTSGVDMPCGADYQTLKLLGLRPSVKRYMYQQGCFAGGTVLRKDLAEN 180  
QY 181 NKGARVLVVCSEVTAVTFRGPDSTHLDLSLVGQALFGDGAALIVGSDPVPPIEIEKPIFEMV 240  
DB 181 NKGARVLVVCSEVTAVTFRGPDSTHLDLSLVGQALFGDGAALIVGSDPVPPIEIEKPIFEMV 240  
QY 241 WTAQTAPDSEGAIDGHLREAGLTTHLLKDVPGIIVSKNITKALVEAFEPGLGSDYNSIFW 300  
DB 241 WTAQTAPDSEGAIDGHLREAGLTTHLLKDVPGIIVSKNITKALVEAFEPGLGSDYNSIFW 300  
QY 301 IAHPGGPAILDQVEOKLALKEPKMNATREVLSEYGNMSSACVLFILDEMRRKKSSTONGLGT 360  
DB 301 IAHPGGPAILDQVEOKLALKEPKMNATREVLSEYGNMSSACVLFILDEMRRKKSSTONGLGT 360  
QY 361 TGEGLWGVLFPGFGLTIETVLRSAI 389  
DB 361 TGEGLWGVLFPGFGLTIETVLRSAI 389

RESULT 15

US-10-424-599-171124  
Sequence 171124, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 171124  
LENGTH: 388  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_12553C.1.pap  
US-10-424-599-171124

Query Match 89.9%; Score 1795.5; DB 15; Length 388;  
Best Local Similarity 87.7%; Pred. No. 7.3e-172; Indels 1; Gaps 1;  
Matches 341; Conservative 26; Mismatches 21; Indels 1; Gaps 1;  
QY 1 MVSVEIRKAQRAEGPATILAIATPANPCVQSTYDPDFYFKITNSEHKTTELKEKFORMC 60  
DB 1 MVSVEIRKAQRAEGPATILAIATPANPCVQSTYDPDFYFKITNSEHKTTELKEKFORMC 60  
QY 61 DKSMIKRYYNLTETILKENPNVCEYMAPSLDARDQMVVVEPRLGKEAAVKAIKEWGQP 120  
DB 61 DKSMIKRYYNLTETILKENPNVCEYMAPSLDARDQMVVVEPRLGKEAAVKAIKEWGQP 120  
QY 121 KSKITHLIVCTTSGVDMPCGADYQTLKLLGLRPSVKRYMYQQGCFAGGTVLRKDLAEN 180  
DB 121 KSKITHLIVCTTSGVDMPCGADYQTLKLLGLRPSVKRYMYQQGCFAGGTVLRKDLAEN 180

DB 121 KSKITHLIVCTTSGVDMPCGADYQTLKLLGLRPSVKRYMYQQGCFAGGTVLRKDLAEN 180  
QY 181 NKGARVLVVCSEVTAVTFRGPDSTHLDLSLVGQALFGDGAALIVGSDPVPPIEIEKPIFEMV 240  
DB 181 NKGARVLVVCSEVTAVTFRGPDSTHLDLSLVGQALFGDGAALIVGSDPVPPIEIEKPIFEMV 240  
QY 241 WTAQTAPDSEGAIDGHLREAGLTTHLLKDVPGIIVSKNITKALVEAFEPGLGSDYNSIFW 300  
DB 241 WTAQTAPDSEGAIDGHLREAGLTTHLLKDVPGIIVSKNITKALVEAFEPGLGSDYNSIFW 300  
QY 301 IAHPGGPAILDQVEOKLALKEPKMNATREVLSEYGNMSSACVLFILDEMRRKKSSTONGLGT 360  
DB 301 IAHPGGPAILDQVEOKLALKEPKMNATREVLSEYGNMSSACVLFILDEMRRKKSSTONGLGT 360  
QY 361 TGEGLWGVLFPGFGLTIETVLRSAI 389  
DB 361 TGEGLWGVLFPGFGLTIETVLRSAI 389

Search completed: April 22, 2005, 21:57:07  
Job time : 138 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2005, 21:27:00 ; Search time 172 Seconds  
(without alignments)  
1158.133 Million cell updates/sec

Title: US-10-031-918A-1  
Perfect score: 1997  
Sequence: 1 MVSVEIRKAQRAEGPATIL.....LFGFGGLTIETVLRVAI 389

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot\_03.\*

1: uniprot\_prot.\*

2: uniprot\_tmbl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1994	99.8	389	1 CHS2 MEDSA	P30074 medicago sa
2	1969	98.6	389	1 CHS4 MEDSA	P30075 medicago sa
3	1966	98.4	389	2 Q9FEW4	Q9few4 medicago tr
4	1941	97.2	389	1 CHS8 MEDSA	P30076 medicago sa
5	1938	97.0	389	1 CHS5 MEDSA	P51078 medicago sa
6	1937	97.0	389	1 CHS5_PEA	Q23884 pisum sativ
7	1932	96.7	389	1 CHS2_PEA	Q01287 pisum sativ
8	1930	96.6	389	1 CHS3_PEA	O23883 pisum sativ
9	1927	96.5	389	1 CHS2 TRISU	P51084 trifolium s
10	1926	96.4	389	1 CHS1_PEA	Q01286 pisum sativ
11	1926	96.4	389	1 CHS9 MEDSA	P30077 medicago sa
12	1925	96.4	389	1 CHS4_PEA	O23882 pisum sativ
13	1920	96.1	389	1 CHS6 TRISU	P51088 trifolium s
14	1915	95.9	389	1 CHS6_PEA	Q01288 pisum sativ
15	1914	95.8	389	1 CHS5 TRISU	P51087 trifolium s
16	1912	95.7	389	1 CHS3 TRISU	P51085 trifolium s
17	1911	95.7	389	1 CHS3 MEDSA	P51077 medicago sa
18	1908	95.5	389	1 CHS1 TRISU	P51083 trifolium s
19	1908	95.5	389	2 Q9ZRV7	Q9zrv7 cicer ariet
20	1903	95.3	389	1 CHS1 CIRCAR	Q9sm14 cicer ariet
21	1881.5	94.2	390	1 CHS7 ONOVI	O22586 onobrychis
22	1880	94.1	389	1 CHS1 MEDSA	P30073 medicago sa
23	1859	93.6	389	2 Q84V87	Q84v87 glycine max
24	1857	93.5	389	1 CHS7_PUELO	P23569 pueraria lo
25	1854	92.8	390	2 Q854Y7	Q854y7 senna alata
26	1844	92.3	389	1 CHS7 PHAVU	P49440 phaseolus v
27	1839	92.1	389	1 CHS7 SOYBN	P30081 glycine max
28	1838	92.0	389	2 Q84XJ8	Q84xa8 arachis hyp
29	1825	91.4	389	2 Q854Y6	Q854y6 senna alata
30	1818	91.0	389	1 CHS4_PEA	P51081 pisum sativ
31	1817	91.0	389	2 Q854Y8	Q854y8 senna alata

RESULT 1

CHS2 MEDSA

ID CHS2 MEDSA STANDARD; PRT; 389 AA.

AC P30074;

DT 01-APR-1993 (Rel. 25, Created)

DT 01-APR-1993 (Rel. 25, Last sequence update)

DT 25-OCT-2004 (Rel. 45, Last annotation update)

DE Chalcone synthase 2 (EC 2.3.1.74) (Naringenin-chalcone synthase 2).

GN Name=CHS2;

OS Medicago sativa (Alfalfa).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC Eurosid1; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.

OX NCBI\_TaxID=3879;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=93283629; PubMed=8507827;

RA Junghans H., Dalkin K., Dixon R.A.;

RT "Stress responses in alfalfa (Medicago sativa L.). 15.

RT Characterization and expression patterns of members of a subset of the

RT chalcone synthase multigene family.";

RL Plant Mol. Biol. 22:239-253(1993).

RN [2]

RP X-RAY CRYSTALLOGRAPHY (1.69 ANGSTROMS).

RX MEDLINE=20120577; PubMed=10653632; DOI=10.1021/bi991489f;

RA Jez J.M., Ferrer J.L., Bowman M.E., Dixon R.A., Noel J.P.;

RT "Dissection of malonyl-coenzyme A decarboxylation from polyketide

RT formation in the reaction mechanism of a plant polyketide synthase.";

RL Biochemistry 39:890-902(2000).

CC -!- FUNCTION: The primary product of this enzyme is 4,2',4',6'-

CC tetrahydroxychalcone (also termed naringenin-chalcone or Chalcone)

CC which can under specific conditions spontaneously isomerize into

CC naringenin.

CC -!- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA +

CC naringenin-chalcone + 3 CO(2).

CC -!- PATHWAY: Part of the biosynthetic pathway for all classes of

CC flavonoids, a large class of secondary plant metabolites, many of

CC which are brightly colored.

CC -!- SIMILARITY: Belongs to the chalcone/stilbene synthases family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; L02902; AAA02824.1; -.

DR PIR; S35164; S35164.

DR PDB; 1BI5; X-ray; A=1-389.

DR PDB; 1BQ6; X-ray; A=2-389.

DR PDB; 1CGK; X-ray; A=1-389.

DR PDB; 1CGZ; X-ray; A=1-389.

P51082 pisum sativ  
P24826 glycine max  
Q6X0N0 glycine max  
Q6X0M9 glycine max  
P48406 glycine max  
P19168 glycine max  
P51089 vigna ungu  
Q816d2 rhododendro  
P17957 glycine max  
P30080 glycine max  
P30081 glycine max  
Q9ZVR8 casuarina g  
Q815n6 malus domes  
Q8W3P6 vitis vinif  
Q9FUB7 hypericum a

ALIGNMENTS

DR PDB; 1CHW; X-ray; A/B=1-389.  
 DR PDB; 1CWL; X-ray; A=1-389.  
 DR PDB; 1SGF; X-ray; A=1-389.  
 DR PDB; 1D6H; X-ray; A=3-389.  
 DR PDB; 1D6I; X-ray; A/B=2-389.  
 DR PDB; 1I86; X-ray; A=1-389.  
 DR PDB; 1I88; X-ray; A/B=1-389.  
 DR PDB; 1I89; X-ray; A/B=1-389.  
 DR PDB; 1I8B; X-ray; A/B=1-389.  
 DR PDB; 1LWX; X-ray; A=1-389.  
 DR InterPro; IPR001099; N-C synthase.  
 DR InterPro; IPR011141; PKS III.  
 DR Pfam; PF02797; Chal sti synt C; 1.  
 DR Pfam; PF00195; Chal sti synt\_N; 1.  
 DR PIRSF; PIRSF000451; PKS\_III; 1.  
 DR ProDom; PD000453; N-C synthase; 1.  
 DR PROSITE; PS00441; CHALCONE SYNTH; 1.  
 KW 3D-structure; Acyltransferase; Flavonoid biosynthesis;  
 KW Multigene family; Transferase.  
 FT ACT SITE 164 164  
 FT HELIX 4 11  
 FT STRAND 18 25  
 FT STRAND 30 32  
 FT TURN 34 35  
 FT HELIX 36 43  
 FT TURN 44 45  
 FT TURN 47 48  
 FT STRAND 50 62  
 FT HELIX 67 69  
 FT STRAND 74 78  
 FT HELIX 79 79  
 FT TURN 81 84  
 FT HELIX 91 117  
 FT HELIX 121 123  
 FT STRAND 126 131  
 FT HELIX 140 148  
 FT TURN 149 149  
 FT TURN 152 153  
 FT STRAND 155 160  
 FT TURN 164 165  
 FT HELIX 166 179  
 FT TURN 180 180  
 FT TURN 182 183  
 FT STRAND 185 192  
 FT HELIX 194 196  
 FT TURN 197 197  
 FT TURN 203 204  
 FT HELIX 206 214  
 FT STRAND 215 215  
 FT STRAND 218 225  
 FT TURN 230 232  
 FT STRAND 237 246  
 FT TURN 249 249  
 FT TURN 251 252  
 FT STRAND 254 255  
 FT STRAND 258 259  
 FT TURN 260 261  
 FT STRAND 262 263  
 FT STRAND 266 267  
 FT HELIX 271 287  
 FT HELIX 288 290  
 FT TURN 291 291  
 FT TURN 295 296  
 FT STRAND 299 302  
 FT HELIX 307 317  
 FT TURN 318 318  
 FT HELIX 321 324  
 FT HELIX 325 334  
 FT STRAND 336 336  
 FT STRAND 338 340  
 FT HELIX 341 355  
 FT TURN 356 357  
 FT STRAND 360 360

FT TURN 361 364  
 FT STRAND 367 374  
 FT TURN 375 377  
 FT STRAND 378 386  
 FT STRAND 388 389  
 SQ SEQUENCE 389 AA; 42706 MW; E03422EE332A5408 CRC64;  
 Query Match 99.8%; Score 1994; DB 1; Length 389;  
 Best Local Similarity 99.7%; Pred. No. 7.7e-136;  
 Matches 388; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MVSVEIRKQRAEGPATILAIAGTANPANCVEQSTYDPDFYFKITNSHKTELKPKFORMC 60  
 DB 1 MVSVEIRKQRAEGPATILAIAGTANPANCVEQSTYDPDFYFKITNSHKTELKPKFORMC 60  
 QY 61 DKSMIKRMYMLTEIILKENPNVCEYMAPSLDARQDMVVVEVPRLGKAAVKAKEWGQP 120  
 DB 61 DKSMIKRMYMLTEIILKENPNVCEYMAPSLDARQDMVVVEVPRLGKAAVKAKEWGQP 120  
 QY 121 KSKITHLIVCTTSGVDMFGADYQTLKLGLRPYVKRYMYQOGXFGAGTVLRLAKDLAEN 180  
 DB 121 KSKITHLIVCTTSGVDMFGADYQTLKLGLRPYVKRYMYQOGXFGAGTVLRLAKDLAEN 180  
 QY 181 NKGARVLVVCSEVTAFTFRGSDTHLSDSLVQALFGDGAALIVGSDPVPEIEKPIFEMV 240  
 DB 181 NKGARVLVVCSEVTAFTFRGSDTHLSDSLVQALFGDGAALIVGSDPVPEIEKPIFEMV 240  
 QY 241 WTAQTAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAPEPLGSDYNSIFW 300  
 DB 241 WTAQTAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAPEPLGSDYNSIFW 300  
 QY 301 IAHPGGPAILDQVEQKLALPEKNATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
 DB 301 IAHPGGPAILDQVEQKLALPEKNATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
 QY 361 TGEGLWGLVFGFGPGGLTIETVLRSAI 389  
 DB 361 TGEGLWGLVFGFGPGGLTIETVLRSAI 389  
 RESULT 2  
 CHS4 MEDSA STANDARD; PRT; 389 AA.  
 ID CHS4 MEDSA  
 AC P30075;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Chalcone synthase 4 (EC 2.3.1.74) (Naringenin-chalcone synthase 4)  
 DE (CHS12-1).  
 GN Name=CHS4;  
 OS Medicago sativa (Alfalfa).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.  
 OX NCBI\_TaxID=3879;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Iroquois; TISSUE=ROOT;  
 RX MEDLINE=94250839; PubMed=8193301;  
 RA McKnann H.I., Hirsch A.M.;  
 RT "Isolation of chalcone synthase and chalcone isomerase cDNAs from  
 alfalfa (Medicago sativa L.): highest transcript levels occur in young  
 roots and root tips.";  
 RL Plant Mol. Biol. 24:767-777(1994).  
 RN [2]  
 RP SEQUENCE OF 7-389 FROM N.A.  
 RX MEDLINE=93283629; PubMed=8507827;  
 RA Jungmans H., Dalkin K., Dixon R.A.;  
 RT "Stress responses in alfalfa (Medicago sativa L.). 15.  
 RT Characterization and expression patterns of members of a subset of the  
 chalcone synthase multigene family.";  
 RL Plant Mol. Biol. 22:239-253(1993).  
 CC !- FUNCTION: The primary product of this enzyme is 4,2',4',6'-

CC tetrahydroxychalcone (also termed naringenin-chalcone or chalcone)  
 CC which can under specific conditions spontaneously isomerize into  
 CC naringenin.  
 CC -1- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +  
 CC naringenin-chalcone + 3 CO(2).  
 CC -1- PATHWAY: Part of the biosynthetic pathway for all classes of  
 CC flavonoids, a large class of secondary plant metabolites, many of  
 CC which are brightly colored.  
 CC -1- DEVELOPMENTAL STAGE: Highest expression in young root tips.  
 CC -1- SIMILARITY: Belongs to the chalcone/stilbene synthases family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U01021; AAB41559.1; -;  
 CC EMBL; L02903; AAA02825.1; ALT\_INIT.  
 CC PIR; S44370; S44370.  
 CC HSSP; P30074; 1CGZ.  
 CC InterPro; IPR001099; N-C\_synthase.  
 CC InterPro; IPR011141; PKS\_III.  
 CC Pfam; PF02797; Chal\_sti\_synt\_C; 1.  
 CC Pfam; PF00195; Chal\_sti\_synt\_N; 1.  
 CC PIRSF; PIRSF000451; PKS\_III; 1.  
 CC ProDom; PD000453; N-C\_synthase; 1.  
 CC PROSITE; PS00441; CHALCONE\_SYNTH; 1.  
 CC Acyltransferase; Flavonoid biosynthesis; Multigene family;  
 CC Transferrase.

KW Transferrase.

FT ACT SITE 164 164 BY similarity.

SQ SEQUENCE 389 AA; 42735 MW; A969AFD7D7483C49 CRC64;

Query Match 98.6%; Score 1969; DB 1; Length 389;

Best Local Similarity 98.5%; Pred. No. 4.9e-134;

Matches 383; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 MVSVEIRKAQRAEGPATILAI GTANPANCVEQSTYDPDFYFKITNSEHKTLEKEFORMC 60  
 DB 1 MVSVEIRKAQRAEGPATILAI GTANPANCVEQSTYDPDFYFKITNSEHKTLEKEFORMC 60  
 QY 61 DKSMIKRRYMYLITEILKENPNVCYMAPSLDARQDMVVVPRLGKEAAVKAKEWGQP 120  
 DB 61 DKSMIKRRYMYLITEILKENPNVCYMAPSLDARQDMVVVPRLGKEAAVKAKEWGQP 120  
 QY 121 KSKITHLIVCTTSGVDMPCADYQLT KLGLRPVVKRYMYQQGXFAGGTVLR LAKDLAEN 180  
 DB 121 KSKITHLIVCTTSGVDMPCADYQLT KLGLRPVVKRYMYQQGXFAGGTVLR LAKDLAEN 180  
 QY 181 NKGARVLVVCSEVTAVTFRGPSDTHLDSL VGQALFGDGAALIVGSDPVPEIEKPIFEV 240  
 DB 181 NKGARVLVVCSEVTAVTFRGPSDTHLDSL VGQALFGDGAALIVGSDPVPEIEKPIFEV 240  
 QY 241 WTAQTIAPDSEGAIDGHLREAGLT FHLKDVPGIVSKNITKALVEAFPELIGISDYSNIFW 300  
 DB 241 WTAQTIAPDSEGAIDGHLREAGLT FHLKDVPGIVSKNITKALVEAFPELIGISDYSNIFW 300  
 QY 301 IAHPGGPAILDQVEQKALPEKRNATREVLSEYGNMSSACVLFILDEMRRKKSSTQGLTKT 360  
 DB 301 IAHPGGPAILDQVEQKALPEKRNATREVLSEYGNMSSACVLFILDEMRRKKSSTQGLTKT 360  
 QY 361 TCGLEWGLVFGFGPGGLTIETVTLRSVAI 389  
 DB 361 TCGLEWGLVFGFGPGGLTIETVTLRSVAI 389

RESULT 3

Q9FEW4 PRELIMINARY; PRT; 389 AA.

ID Q9FEW4

AC Q9FEW4

DT 01-MAR-2001 (TrEMBLrel. 16, Created)

DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Putative chalcone synthase.  
 GN Name=chcl;  
 OS Medicago truncatula (Barrel medic)  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.  
 ON NCBI\_TaxID=3880;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RA Bonanomi A., Oetiker J.H., Guggenheim R., Boller T., Andress W.,  
 RA Voegel-Lange R.;  
 RT "Arbuscular mycorrhiza in mini-mycorrhizotrons: first contact of  
 RT Medicago truncatula roots with Glomus intraradices induces Chalcone  
 RT synthase.";  
 RL New Phytol. 150:573-582(2001).  
 CC -1- SIMILARITY: Belongs to the chalcone/stilbene synthases family.  
 DR HSSP; P30074; 1B15.  
 DR EMBL; AJ277211; CAC20725.1; -;  
 DR GO; GO:0008415; F:acyltransferase activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0009058; P:biosynthesis; IEA.  
 DR InterPro; IPR001099; N-C\_synthase.  
 DR InterPro; IPR011141; PKS\_III.  
 DR Pfam; PF02797; Chal\_sti\_synt\_C; 1.  
 DR PIRSF; PIRSF000451; PKS\_III; 1.  
 DR ProDom; PD000453; N-C\_synthase; 1.  
 DR PROSITE; PS00441; CHALCONE\_SYNTH; 1.  
 DR Acyltransferase; Transferrase.

KW Acyltransferase; Transferrase.

SQ SEQUENCE 389 AA; 42802 MW; 91EC6FC919DFE7AE CRC64;

Query Match 98.4%; Score 1966; DB 2; Length 389;

Best Local Similarity 98.2%; Pred. No. 8.1e-134;

Matches 382; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 MVSVEIRKAQRAEGPATILAI GTANPANCVEQSTYDPDFYFKITNSEHKTLEKEFORMC 60  
 DB 1 MVSVEIRKAQRAEGPATILAI GTANPANCVEQSTYDPDFYFKITNSEHKTLEKEFORMC 60  
 QY 61 DKSMIKRRYMYLITEILKENPNVCYMAPSLDARQDMVVVPRLGKEAAVKAKEWGQP 120  
 DB 61 DKSMIKRRYMYLITEILKENPNVCYMAPSLDARQDMVVVPRLGKEAAVKAKEWGQP 120  
 QY 121 KSKITHLIVCTTSGVDMPCADYQLT KLGLRPVVKRYMYQQGXFAGGTVLR LAKDLAEN 180  
 DB 121 KSKITHLIVCTTSGVDMPCADYQLT KLGLRPVVKRYMYQQGXFAGGTVLR LAKDLAEN 180  
 QY 181 NKGARVLVVCSEVTAVTFRGPSDTHLDSL VGQALFGDGAALIVGSDPVPEIEKPIFEV 240  
 DB 181 NKGARVLVVCSEVTAVTFRGPSDTHLDSL VGQALFGDGAALIVGSDPVPEIEKPIFEV 240  
 QY 241 WTAQTIAPDSEGAIDGHLREAGLT FHLKDVPGIVSKNITKALVEAFPELIGISDYSNIFW 300  
 DB 241 WTAQTIAPDSEGAIDGHLREAGLT FHLKDVPGIVSKNITKALVEAFPELIGISDYSNIFW 300  
 QY 301 IAHPGGPAILDQVEQKALPEKRNATREVLSEYGNMSSACVLFILDEMRRKKSSTQGLTKT 360  
 DB 301 IAHPGGPAILDQVEQKALPEKRNATREVLSEYGNMSSACVLFILDEMRRKKSSTQGLTKT 360  
 QY 361 TCGLEWGLVFGFGPGGLTIETVTLRSVAI 389  
 DB 361 TCGLEWGLVFGFGPGGLTIETVTLRSVAI 389

RESULT 4

CHS8 MEDSA

ID CHS8 MEDSA STANDARD; PRT; 389 AA.

AC P30076;

DT 01-APR-1993 (Rel. 25, Created)

DT 01-APR-1993 (Rel. 25, Last sequence update)

DT 25-OCT-2004 (Rel. 45, Last annotation update)

DE Chalcone synthase 8 (EC 2.3.1.74) (Naringenin-chalcone synthase 8).

GN Name=CHS8;  
OS Medicago sativa (Alfalfa).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.  
OX NCBI\_TaxID=3879;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93283629; PubMed=8507827;  
RA Jungmans H., Dalkin K., Dixon R.A.;  
RT "Stress responses in alfalfa (Medicago sativa L.). 15.  
RT Characterization and expression patterns of members of a subset of the  
RT chalcone synthase multigene family.";  
RL Plant Mol. Biol. 22:239-253(1993).  
CC -!- FUNCTION: The primary product of this enzyme is 4,2',4',6'-  
CC tetrahydroxylchalcone (also termed naringenin-chalcone or chalcone)  
CC which can under specific conditions spontaneously isomerize into  
CC naringenin.  
CC -!- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +  
CC naringenin-chalcone + 3 CO(2).  
CC -!- PATHWAY: Part of the biosynthetic pathway for all classes of  
CC flavonoids, a large class of secondary plant metabolites, many of  
CC which are brightly colored.  
CC -!- SIMILARITY: Belongs to the chalcone/stilbene synthases family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; J02904; AAA02826.1; -.  
CC PIR; S35166; S35166.  
CC HSSP; P30074; 1CGZ.  
CC InterPro; IPR001099; N-C synthase.  
CC Pfam; PF02797; Chal\_sti\_synt\_C; 1.  
CC Pfam; PF00195; Chal\_sti\_synt\_N; 1.  
CC PIRSF; PIRSF000451; PKS\_III; 1.  
CC ProDom; PD000453; N-C synthase; 1.  
CC PROSITE; PS00441; CHALCONE SYNTH; 1.  
KW Acyltransferase; Flavonoid biosynthesis; Multigene family;  
KW Transferase. 164 164 By similarity.  
SQ ACT\_SITE 389 AA; 42549 MW; ASDFC41B4A9DBF2C CRC64;  
SEQUENCE 389 AA; 42549 MW; ASDFC41B4A9DBF2C CRC64;  
Query Match 97.2%; Score 1941; DB 1; Length 389;  
Best Local Similarity 97.4%; Pred. No. 5.2e-132;  
Matches 379; Conservative 3; Mismatches 7; Indels 0; Gaps 0;  
QY 1 MVSVEIRKQAEAGPATILAI GTANPANCVEQSTYDPDFYFKITNSEHKTLEKPFQRC 60  
DB 1 MVSVEIRKQAEAGPATILAI GTANPANCVEQSTYDPDFYFKITNSEHKTLEKPFQRC 60  
QY 61 DKSMIKRRYMYLTEELTKENPNVCEYMAPSLDARDMVVVEPRGLCKEAAVKAIEWGQP 120  
DB 61 DKSMIKRRYMYLTEELTKENPNVCEYMAPSLDARDMVVVEPRGLCKEAAVKAIEWGQP 120  
QY 121 KSKITHLIVCTTSGVDMPGADYQLTKLGLRPPVKKRYMYQQQXFGAGTGLRLAKDLAEN 180  
DB 121 KSKITHLIVCTTSGVDMPGADYQLTKLGLRPPVKKRYMYQQQXFGAGTGLRLAKDLAEN 180  
QY 181 NKGARVLVVCSEVTA VTFRGPSTHLDLSLVGALFGDGAALIVGSDPVPPEIEKPIFENV 240  
DB 181 NKGARVLVVCSEVTA VTFRGPSTHLDLSLVGALFGDGAALIVGSDPVPPEIEKPIFENV 240  
QY 241 WTAQTTAPSEGAIDGHLREAGLTFFHLKDVPGIVSKNITKALVEAFELGSDVNSIPW 300  
DB 241 WTAQTTAPSEGAIDGHLREAGLTFFHLKDVPGIVSKNITKALVEAFELGSDVNSIPW 300  
QY 301 IAHFGGPAILDQVEQKLALPKPKMKAATREVLSEYGNMSSACVLVILDEMRRKKSADGLKT 360

Db 301 IAHFGGPAILDQVEQKLALPKPKMKAATREVLSEYGNMSSACVLVILDEMRRKKSADGLKT 360  
QY 361 TGEGLGVNGLFGPGPGGLTIETVLRSAI 389  
DB 361 TGEGLGVNGLFGPGPGGLTIETVLRSAI 389  
RESULT 5  
CHS5 MEDSA STANDARD; PRT; 389 AA.  
AC PS1078;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Chalcone synthase 4-2 (EC 2.3.1.74) (Naringenin-chalcone synthase 4-2).  
DE 2).  
GN Name=CHS4-2; Synonyms=CHS1;  
OS Medicago sativa (Alfalfa).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.  
OX NCBI\_TaxID=3879;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=cv. Nagvazenas; TISSUE=Leaf;  
RX MEDLINE=94100583; PubMed=8274775;  
RA Esnault R., Buffard D., Breda C., Sallaud C., El-Turk J.,  
RA Kondorosi A.;  
RT "Pathological and molecular characterizations of alfalfa interactions  
RT with compatible and incompatible bacteria, Xanthomonas campestris pv.  
RT alfalfae and Pseudomonas syringae pv. pisi.";  
RL Mol. Plant Microbe Interact. 6:655-664(1993).  
RN [2]  
RP SEQUENCE OF 19-389 FROM N.A.  
RX STRAIN=cv. Iroquois; TISSUE=Root nodules;  
RX MEDLINE=94250839; PubMed=8193301;  
RA McKhann H.I., Hirsch A.M.;  
RT "Isolation of chalcone synthase and chalcone isomerase cDNAs from  
RT alfalfa (Medicago sativa L.): highest transcript levels occur in young  
RT roots and root tips.";  
RL Plant Mol. Biol. 24:767-777(1994).  
CC -!- FUNCTION: The primary product of this enzyme is 4,2',4',6'-  
CC tetrahydroxylchalcone (also termed naringenin-chalcone or chalcone)  
CC which can under specific conditions spontaneously isomerize into  
CC naringenin.  
CC -!- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +  
CC naringenin-chalcone + 3 CO(2).  
CC -!- PATHWAY: Part of the biosynthetic pathway for all classes of  
CC flavonoids, a large class of secondary plant metabolites, many of  
CC which are brightly colored.  
CC -!- DEVELOPMENTAL STAGE: Highest expression in young root tips.  
CC -!- INDUCTION: By infection.  
CC -!- SIMILARITY: Belongs to the chalcone/stilbene synthases family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X68106; CAA48226.1; -.  
CC PIR; S26414; S26414.  
CC HSSP; P30074; 1BI5.  
CC InterPro; IPR001099; N-C synthase.  
CC Pfam; PF02797; Chal\_sti\_synt\_C; 1.  
CC Pfam; PF00195; Chal\_sti\_synt\_N; 1.  
CC PIRSF; PIRSF000451; PKS\_III; 1.  
CC ProDom; PD000453; N-C synthase; 1.

DR PROSITE; PS00441; CHALCONE SYNTH; 1.  
 KW Acyltransferase; Flavonoid biosynthesis; Multigene family;  
 FT ACT SITE 164 By similarity.  
 FT CONFLICT 86 I -> Y (in Ref. 2).  
 FT CONFLICT 94 W -> R (in Ref. 2).  
 SQ SEQUENCE 389 AA; 42661 MW; B33674EC7709F170 CRC64;  
 Query Match 97.0%; Score 1938; DB 1; Length 389;  
 Best Local Similarity 97.2%; Pred. No. 8.6e-132;  
 Matches 378; Conservative 2; Mismatches 9; Indels 0; Gaps 0;  
 QY 1 MVSVSIRKAQRAEGPATILAI GTANPANCVEQSTYDPFYKLTNSEHKTTELKQFQRC 60  
 DB 1 MVSVSIRKAQRAEGPATILAI GTANPANCVEQSTYDPFYKLTNSEHKTTELKQFQRC 60  
 QY 61 DKSMIKRRYMYLITEELKENPNVCYMAPSLDARDQVNVVPRGLGKEAAVKAKEWGOP 120  
 DB 61 DKSMIKRRYMYLITEELKENPNVCYMAPSLDARDQVNVVPRGLGKEAAVKAKEWGOP 120  
 QY 121 KSKITHLIVCTTSGVDMPCADYQTLKGLRPPVYKRYMYQQXFGAGTVLRLAKDLAEN 180  
 DB 121 KSKITHLIVCTTSGVDMPCADYQTLKGLRPPVYKRYMYQQXFGAGTVLRLAKDLAEN 180  
 QY 181 NKGARVLVVCSEVTATFRGSDTHLDSLVGQALFGDGAALIVGSDPVEIEKPIFEMV 240  
 DB 181 NKGARVLVVCSEVTATFRGSDTHLDSLVGQALFGDGAALIVGSDPVEIEKPIFEMV 240  
 QY 241 WTAQTIAPDSEGAIDGHLREAGLTFLHLLKDVPGIVSKNITKALVEAFELGSDYNSIFW 300  
 DB 241 WTAQTIAPDSEGAIDGHLREAGLTFLHLLKDVPGIVSKNITKALVEAFELGSDYNSIFW 300  
 QY 301 IAHPGGPAILDQVEQKALKEPKKATREVLSEYGNMSSACVLFILDEMRRKSTQNGLKT 360  
 DB 301 IAHPGGPAILDQVEQKALKEPKKATREVLSEYGNMSSACVLFILDEMRRKSTQNGLKT 360  
 QY 361 TGEGLWGLVFGPGGLTIETVLRSAI 389  
 DB 361 TGEGLWGLVFGPGGLTIETVLRSAI 389  
 RESULT 6  
 CHS5\_PEA STANDARD; PRT; 389 AA.  
 AC Q23884;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Chalcone synthase 5 (EC 2.3.1.74) (Naregenin-chalcone synthase 5).  
 GN Name=CHS5;  
 OS Pisum sativum (Garden pea).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.  
 OX NCBI\_TaxID=3888;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Midoriusui;  
 RX MEDLINE=97374447; PubMed=9230896; DOI=10.1007/s004380050471;  
 RA Ito M., Ichinohe Y., Kato H., Shiraiishi T., Yamada T.;  
 RT "Molecular evolution and functional relevance of the chalcone synthase genes of pea."  
 RL Mol. Gen. Genet. 255:28-37(1997).  
 CC -!- FUNCTION: The primary product of this enzyme is 4,2',4',6'-tetrahydroxychalcone (also termed naringenin-chalcone or chalcone) which can under specific conditions spontaneously isomerize into naringenin.  
 CC -!- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA + naringenin-chalcone + 3 CO(2).  
 CC -!- PATHWAY: Part of the biosynthetic pathway for all classes of flavonoids, a large class of secondary plant metabolites, many of which are brightly colored.  
 CC -!- SIMILARITY: Belongs to the chalcone/stilbene synthases family.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; D88262; BAA22044.1; -.  
 HSSP; P30074; ICGZ.  
 InterPro; IPR01099; N-C synthase.  
 InterPro; IPR01141; PKS III.  
 Pfam; PF02797; Chal\_sti\_synt\_C; 1.  
 Pfam; PF00195; Chal\_sti\_synt\_N; 1.  
 PIRSF; PIRSF000451; PKS\_III; 1.  
 ProDom; PD000453; N-C synthase; 1.  
 PROSITE; PS00441; CHALCONE SYNTH; 1.  
 Acyltransferase; Flavonoid biosynthesis; Multigene family;  
 KW Transferrase.  
 FT ACT SITE 164 164 By similarity.  
 SQ SEQUENCE 389 AA; 42816 MW; 32272DC9B45DEA CRC64;  
 Query Match 97.0%; Score 1937; DB 1; Length 389;  
 Best Local Similarity 96.4%; Pred. No. 1e-131;  
 Matches 375; Conservative 6; Mismatches 8; Indels 0; Gaps 0;  
 QY 1 MVSVSIRKAQRAEGPATILAI GTANPANCVEQSTYDPFYKLTNSEHKTTELKQFQRC 60  
 DB 1 MVSVSIRKAQRAEGPATILAI GTANPANCVEQSTYDPFYKLTNSEHKTTELKQFQRC 60  
 QY 61 DKSMIKRRYMYLITEELKENPNVCYMAPSLDARDQVNVVPRGLGKEAAVKAKEWGOP 120  
 DB 61 DKSMIKRRYMYLITEELKENPNVCYMAPSLDARDQVNVVPRGLGKEAAVKAKEWGOP 120  
 QY 121 KSKITHLIVCTTSGVDMPCADYQTLKGLRPPVYKRYMYQQXFGAGTVLRLAKDLAEN 180  
 DB 121 KSKITHLIVCTTSGVDMPCADYQTLKGLRPPVYKRYMYQQXFGAGTVLRLAKDLAEN 180  
 QY 181 NKGARVLVVCSEVTATFRGSDTHLDSLVGQALFGDGAALIVGSDPVEIEKPIFEMV 240  
 DB 181 NKGARVLVVCSEVTATFRGSDTHLDSLVGQALFGDGAALIVGSDPVEIEKPIFEMV 240  
 QY 241 WTAQTIAPDSEGAIDGHLREAGLTFLHLLKDVPGIVSKNITKALVEAFELGSDYNSIFW 300  
 DB 241 WTAQTIAPDSEGAIDGHLREAGLTFLHLLKDVPAIVSKNIDKALVEAFQPLGSDYNSIFW 300  
 QY 301 IAHPGGPAILDQVEQKALKEPKKATREVLSEYGNMSSACVLFILDEMRRKSTQNGLKT 360  
 DB 301 IAHPGGPAILDQVEQKALKEPKKATREVLSEYGNMSSACVLFILDEMRRKSTQNGLKT 360  
 QY 361 TGEGLWGLVFGPGGLTIETVLRSAI 389  
 DB 361 TGEGLWGLVFGPGGLTIETVLRSAI 389  
 RESULT 7  
 CHS2\_PEA STANDARD; PRT; 389 AA.  
 AC Q01287;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Chalcone synthase 2 (EC 2.3.1.74) (Naregenin-chalcone synthase 2).  
 GN Name=CHS2;  
 OS Pisum sativum (Garden pea).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.  
 OX NCBI\_TaxID=3888;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Midoriusui; TISSUE=Epicotyl;

```

RX MEDLINE=92256801; PubMed=1581561;
RA Ichinose Y., Kawanata S., Yamada T., An C., Kajiura T., Shiraishi T.,
RA Oku H.;
RT "Molecular cloning of chalcone synthase cDNAs from Pisum sativum.";
RL Plant Mol. Biol. 18:1009-1012(1992).
RN [2]
RP SEQUENCE OF 1-128 FROM N.A.
RC STRAIN=cv. Midoriisui; TISSUE=Epicotyl;
RX MEDLINE=93322478; PubMed=8467077;
RA An C., Ichinose Y., Yamada T., Tanaka Y., Shiraishi T., Oku H.;
RT "Organization of the genes encoding chalcone synthase in Pisum
RT sativum.";
RL Plant Mol. Biol. 21:789-803(1993).
CC -!- FUNCTION: The primary product of this enzyme is 4,2',4',6'-
CC tetrahydrochalcone (also termed naringenin-chalcone or chalcone)
CC which can under specific conditions spontaneously isomerize into
CC naringenin.
CC -!- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +
CC naringenin-chalcone + 3 CO(2).
CC -!- PATHWAY: Part of the biosynthetic pathway for all classes of
CC flavonoids, a large class of secondary plant metabolites, many of
CC which are brightly colored.
CC -!- SIMILARITY: Belongs to the chalcone/stilbene synthases family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X63334; CAA44934.1; -.
DR EMBL; D10662; BAA01513.1; -.
DR PIR; S20932; S20932.
DR HSSP; P30074; 1CG2.
DR InterPro; IPR001099; N-C synthase.
DR InterPro; IPR011141; PKS_III.
DR Pfam; PF02797; Chal_sti_synt_C; 1.
DR Pfam; PF00195; Chal_sti_synt_N; 1.
DR PIRSF; PIRSF000451; PKS_III; 1.
DR ProDom; PD000453; N-C synthase; 1.
DR PROSITE; PS00441; CHALCONE SYNTH; 1.
KW Acyltransferase; Flavonoid biosynthesis; Multigene family;
FT ACT SITE 164 164 By similarity.
SQ SEQUENCE 389 AA; 42844 MW; 0585EC37CA4F072 CRC64;

Query Match 96.7%; Score 1932; DB 1; Length 389;
Best Local Similarity 96.4%; Pred. No. 2.3e-131;
Matches 375; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 MVSVEIRKAQRAEGPATILAI GTANPANCVEQSTYDPDFYFKITNSEHKTTELKQFQRC 60
DB 1 MVTVSIRKAQRAEGPATILAI GTANPANCVEQSTYDPDFYFKITNSEHKTTELKQFQRC 60
QY 61 DKSMIKRRYMYLTEETILKENPNVCYMAPSLDARDQMVVVEPRLGKEAAVKAKEWGOP 120
DB 61 DKSMIKRRYMYLTEETILKENPNVCYMAPSLDARDQMVVVEPRLGKEAAVKAKEWGOP 120
QY 121 KSKITHLIVCTTSGVDMPGADYQTLKGLRLPVYKRYMMYQQCGFAGGTVFLRLAKDLAEN 180
DB 121 KSKITHLIVCTTSGVDMPGADYQTLKGLRLPVYKRYMMYQQCGFAGGTVFLRLAKDLAEN 180
QY 181 NKGARVLVVCSEVTAVTFRGPSTHLDLSLVQALFGDGAALIVGSDPVEIEKPIFENV 240
DB 181 NKNARVLVVCSEVTAVTFRGPSTHLDLSLVQALFGDGAALIVGSDPVEIEKPIFENV 240
QY 241 WTAQTAPDSEGAIDCHLREAGITFHLKDVPGIVSKNITKALVEAFELGSDYNSIFW 300
DB 241 WTAQTAPDSEGAIDCHLREAGITFHLKDVPGIVSKNITKALVEAFELGSDYNSIFW 300
QY 301 IAHFGGPAILDQVEOKLALKPEKONATREVLSEYGNMSSACVLFILDEMRRKSTQGLNT 360

```

```

DB 301 IAHFGGPAILDQVEOKLALKPEKONATREVLSEYGNMSSACVLFILDEMRRKSTQGLNT 360
QY 361 TGEGLWGVLFPGFPGFTIETVLRVSVAI 389
DB 361 TGEGLWGVLFPGFPGFTIETVLRVSVAI 389

RESULT 8
CHS3_PEA
ID CHS3_PEA STANDARD; PRT; 389 AA.
AC Q23883;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Chalcone synthase 3 (EC 2.3.1.74) (Naringenin-chalcone synthase 3).
GN Name=CHS3;
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.
OX NCBI_taxid=3888;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Midoriisui;
RX MEDLINE=97374447; PubMed=9230896; DOI=10.1007/s004380050471;
RA Ito M., Ichinose Y., Kato H., Shiraishi T., Yamada T.;
RT "Molecular evolution and functional relevance of the chalcone synthase
RT genes of pea.";
RL Mol. Gen. Genet. 255:28-37(1997).
CC -!- FUNCTION: The primary product of this enzyme is 4,2',4',6'-
CC tetrahydrochalcone (also termed naringenin-chalcone or chalcone)
CC which can under specific conditions spontaneously isomerize into
CC naringenin.
CC -!- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +
CC naringenin-chalcone + 3 CO(2).
CC -!- PATHWAY: Part of the biosynthetic pathway for all classes of
CC flavonoids, a large class of secondary plant metabolites, many of
CC which are brightly colored.
CC -!- SIMILARITY: Belongs to the chalcone/stilbene synthases family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D88261; BAA22043.1; -.
DR HSSP; P30074; 1BI5.
DR InterPro; IPR001099; N-C synthase.
DR InterPro; IPR011141; PKS_III.
DR Pfam; PF02797; Chal_sti_synt_C; 1.
DR Pfam; PF00195; Chal_sti_synt_N; 1.
DR PIRSF; PIRSF000451; PKS_III; 1.
DR ProDom; PD000453; N-C synthase; 1.
DR PROSITE; PS00441; CHALCONE SYNTH; 1.
KW Acyltransferase; Flavonoid biosynthesis; Multigene family;
FT ACT SITE 164 164 By similarity.
SQ SEQUENCE 389 AA; 42802 MW; 70720C8BCC15CAB8 CRC64;

Query Match 96.6%; Score 1930; DB 1; Length 389;
Best Local Similarity 96.1%; Pred. No. 3.3e-131;
Matches 374; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 1 MVSVEIRKAQRAEGPATILAI GTANPANCVEQSTYDPDFYFKITNSEHKTTELKQFQRC 60
DB 1 MVSVEIRKAQRAEGPATILAI GTANPANCVEQSTYDPDFYFKITNSEHKTTELKQFQRC 60
QY 61 DKSMIKRRYMYLTEETILKENPNVCYMAPSLDARDQMVVVEPRLGKEAAVKAKEWGOP 120

```



Db 61 DKSMINRRMYLITEILKENPSCVCEYMAPSLDARQDMVVVPRRLGKEAAVKAKEWGQP 120  
 QY 121 KSKITHLIVCTTSGVDMPGADYQTLKGLRLPVYKRYMYQQXFAAGTGLRLAKDLAEN 180  
 Db 121 KSKITHLIFCTTSGVDMPGADYQTLKGLRLPVYKRYMYQQXFAAGTGLRLAKDLAEN 180  
 QY 181 NKGARVLVVCSEVTATFRGSDTHLDSLVGQALFGDGAALIVGSDPVPTEKPIFEV 240  
 Db 181 NKGARVLVVCSEVTATFRGSDTHLDSLVGQALFGDGAALIVGSDPVPTEKPIFEV 240  
 QY 241 WTAQTIAPDSEGAIDGHLREAGLTFLHLLKDVPGIVSKNITKALVEAFELGSDYNSIFW 300  
 Db 241 WTAQTIAPDSEGAIDGHLREAGLTFLHLLKDVPAIVSKNIDKALVEAFELGSDYNSIFW 300  
 QY 301 IAHPGGPAILDQVEOKLALPKPKMATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
 Db 301 IAHPGGPAILDQVEOKLALPKPKMATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
 QY 361 TGEGLWGLVFGFGPLGTIETVVLRSVAI 389  
 Db 361 TGEGLWGLVFGFGPLGTIETVVLRSVAI 389

RESULT 9

CHS2\_TRISU STANDARD; PRT; 389 AA.  
 ID CHS2\_TRISU STANDARD; PRT; 389 AA.  
 AC P51084;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Chalcone synthase 2 (EC 2.3.1.74) (Naringenin-chalcone synthase 2).  
 GN Names:CHS2;  
 OS Trifolium subterraneum (Subterranean clover).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosid I; Fabales; Fabaceae; Papilionoideae; Trifoliales; Trifolium.  
 OX NCBI\_TaxID=3900;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Karridale; TISSUE=Leaf, and Stem;  
 RX MEDLINE=94171060; PubMed=8125320; DOI=10.1016/0378-1119(94)90785-4;  
 RA Arioli T., Howles P.A., Weimann J.J., Rolfe B.G.;  
 RT "In Trifolium subterraneum, chalcone synthase is encoded by a  
 RT multigene family";  
 RL Gene 138:79-86(1994).  
 CC -!- FUNCTION: The primary product of this enzyme is 4,2',4',6'-  
 CC tetrahydroxychalcone (also termed naringenin-chalcone or chalcone)  
 CC which can under specific conditions spontaneously isomerize into  
 CC naringenin.  
 CC -!- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +  
 CC naringenin-chalcone + 3 CO(2).  
 CC -!- PATHWAY: Part of the biosynthetic pathway for all classes of  
 CC flavonoids, a large class of secondary plant metabolites, many of  
 CC which are brightly colored.  
 CC -!- SIMILARITY: Belongs to the chalcone/stilbene synthases family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC EMBL; M91194; AAA18177.1; -;  
 CC HSP; P30074; IBI5.  
 CC InterPro; IPR010399; N-C synthase.  
 CC InterPro; IPR011141; PKS III.  
 CC Pfam; PF02797; Chal\_sti\_synt\_C; 1.  
 CC Pfam; PF00195; Chal\_sti\_synt\_N; 1.  
 CC FIRSF; FIRSF00451; PKS\_III; 1.  
 CC ProDom; PD000453; N-C synthase; 1.  
 CC PROSITE; PS00441; CHALCONE\_SYNTH; 1.

KW Acyltransferase; Flavonoid biosynthesis; Multigene family;  
 KW Transferase.  
 FT ACT SITE 164 164 By similarity.  
 SQ SEQUENCE 389 AA; 42728 MW; 51200DE3B9C58404 CRC64;  
 Query Match 96.5%; Score 1927; DB 1; Length 389;  
 Best Local Similarity 96.7%; Pred. No. 5.4e-131;  
 Matches 376; Conservative 3; Mismatches 10; Indels 0; Gaps 0;  
 QY 1 MVSVEIRKAQRAEGPATILAICTANPANCVSQSTVPDFYFKITNSEHKTELKEKEQRM 60  
 Db 1 MVSVEIRKAQRAEGPATILAICTANPANCVSQSTVPDFYFKITNSEHKTELKEKEQRM 60  
 QY 61 DKSMIKRMYLITEILKENPSCVCEYMAPSLDARQDMVVVPRRLGKEAAVKAKEWGQP 120  
 Db 61 DKSMIKRMYLITEILKENPSCVCEYMAPSLDARQDMVVVPRRLGKEAAVKAKEWGQP 120  
 QY 121 KSKITHLIVCTTSGVDMPGADYQTLKGLRLPVYKRYMYQQXFAAGTGLRLAKDLAEN 180  
 Db 121 KSKITHLIFCTTSGVDMPGADYQTLKGLRLPVYKRYMYQQXFAAGTGLRLAKDLAEN 180  
 QY 181 NKGARVLVVCSEVTATFRGSDTHLDSLVGQALFGDGAALIVGSDPVPTEKPIFEV 240  
 Db 181 NKGARVLVVCSEVTATFRGSDTHLDSLVGQALFGDGAALIVGSDPVPTEKPIFEV 240  
 QY 241 WTAQTIAPDSEGAIDGHLREAGLTFLHLLKDVPGIVSKNITKALVEAFELGSDYNSIFW 300  
 Db 241 WTAQTIAPDSEGAIDGHLREAGLTFLHLLKDVPGIVSKNIDKALVEAFELGSDYNSIFW 300  
 QY 301 IAHPGGPAILDQVEOKLALPKPKMATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
 Db 301 IAHPGGPAILDQVEOKLALPKPKMATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
 QY 361 TGEGLWGLVFGFGPLGTIETVVLRSVAI 389  
 Db 361 TGEGLWGLVFGFGPLGTIETVVLRSVAI 389

RESULT 10

CHS1\_PEA  
 ID CHS1\_PEA STANDARD; PRT; 389 AA.  
 AC Q01286;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Chalcone synthase 1 (EC 2.3.1.74) (Naringenin-chalcone synthase 1).  
 GN Name:CHS1;  
 OS Pisum sativum (Garden pea).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosid I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.  
 OX NCBI\_TaxID=3888;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Midoriisui; TISSUE=Epicotyl;  
 RX MEDLINE=92256801; PubMed=1581561;  
 RA Ichinose Y., Kawamata S., Yamada T., An C., Kajiwa T., Shiraishi T.,  
 RA Oku H.;  
 RT "Molecular cloning of chalcone synthase cDNAs from Pisum sativum";  
 RL Plant Mol. Biol. 18:1009-1012(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93222478; PubMed=8467077;  
 RA An C., Ichinose Y., Yamada T., Tanaka Y., Shiraishi T., Oku H.;  
 RT "Organization of the genes encoding chalcone synthase in Pisum  
 RT sativum";  
 RL Plant Mol. Biol. 21:789-803(1993).  
 CC -!- FUNCTION: The primary product of this enzyme is 4,2',4',6'-  
 CC tetrahydroxychalcone (also termed naringenin-chalcone or chalcone)  
 CC which can under specific conditions spontaneously isomerize into  
 CC naringenin.  
 CC -!- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +  
 CC naringenin-chalcone + 3 CO(2).

CC -!- PATHWAY: Part of the biosynthetic pathway for all classes of  
 CC flavonoids, a large class of secondary plant metabolites, many of  
 CC which are brightly colored.  
 CC -!- SIMILARITY: Belongs to the chalcone/stilbene synthases family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X63333; CAA44933.1; -;  
 CC EMBL; D10661; BAA01512.1; -;  
 CC PIR; S33610; S33610.  
 CC HSP; P30074; IB15.  
 CC InterPro; IPR001099; N-C synthase.  
 CC InterPro; IPR011141; PKS\_III.  
 CC Pfam; PF02797; Chal\_sti\_synt\_C; 1.  
 CC Pfam; PF00195; Chal\_sti\_synt\_N; 1.  
 CC PIRSF; PIRSF000451; PKS\_III; 1.  
 CC ProDom; PD000453; N-C synthase; 1.  
 CC PROSITE; PS00441; CHALCONE\_SYNTH; 1.  
 CC Acyltransferase; Flavonoid biosynthesis; Multigene family;  
 KW Transferrase.  
 FT ACT SITE 164 164 By similarity.  
 FT SEQUENCE 389 AA; 42802 MW; BEE71F08B0AEB4 CRC64;  
 SQ  
 Query Match 96.4%; Score 1926; DB 1; Length 389;  
 Best Local Similarity 96.1%; Pred. No. 6.3e-131;  
 Matches 374; Conservative 5; Mismatches 10; Indels 0; Gaps 0;  
 QY 1 MVSVEIRKAORAEAGPATILATGTANPANCVEQSTYDPDFYFKITNSEHKTTELKKEKFORMC 60  
 DB 1 MVSVEIRKAORAEAGPATILATGTANPANCVEQSTYDPDFYFKITNSEHKTTELKKEKFORMC 60  
 QY 61 DKSIMKRRYMYLITEILKENPFCVCEYMAPSLDARDQMDVWVVEVPRLGKEAAVKAKEWGQP 120  
 DB 61 DKSIMKRRYMYLITEILKENPFCVCEYMAPSLDARDQMDVWVVEVPRLGKEAAVKAKEWGQP 120  
 QY 121 KSKITHLIVCTTSGVDMPGADYQTLTKLGLRPYKRYMYQOQXCFAGGTVRLAKDLAEN 180  
 DB 121 KSKITHLIFCTTSGVDMPGADYQTLTKLGLRPYKRYMYQOQXCFAGGTVRLAKDLAEN 180  
 QY 181 NKGARVLVVCSEVTATVFRGSDTHLDSLVGOALFGDGAALIVGSDPVPPIEKPIFEMV 240  
 DB 181 NKGARVLVVCSEVTATVFRGSDTHLDSLVGOALFGDGAALIVGSDPVPPIEKPIFEMV 240  
 QY 241 WTAQTIAPDSEGAIDGHLREAGLTFLHLLKDVPGIVSKNITKALVEAFELGIDSYNSIFW 300  
 DB 241 WTAQTIAPDSEGAIDGHLREAGLTFLHLLKDVPGIVSKNIDKALVEAFELGIDSYNSIFW 300  
 QY 301 IAHGPGPAILDQVEQKALKEPKMNAATREVLSEYGNMSSACVLFILDEMRRKSTQNGLKT 360  
 DB 301 IAHGPGPAILDQVEQKALKEPKMNAATREVLSEYGNMSSACVLFILDEMRRKSTQNGLKT 360  
 QY 361 TGEGLWGVLFPGPGGLTETVLRSAI 389  
 DB 361 TGEGLWGVLFPGPGGLTETVLRSAI 389  
 RESULT 11  
 CHS9 MEDSA STANDARD; PRT; 389 AA.  
 ID CHS9 MEDSA  
 AC P30077;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Chalcone synthase 9 (EC 2.3.1.74) (Naringenin-chalcone synthase 9).  
 GN Name=CHS9;  
 OS Medicago sativa (Alfalfa).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC euraoids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae; Medicago.  
 OC NCBI\_TaxID=3879;  
 OC RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93283629; PubMed=8507827;  
 RA Junghans H., Dalkin K., Dixon R.A.;  
 RT "Stress responses in alfalfa (Medicago sativa L.). 15.  
 RT Characterization and expression patterns of members of a subset of the  
 RT chalcone synthase multigene family.";  
 RL Plant Mol. Biol. 22:239-253(1993).  
 CC -!- FUNCTION: The primary product of this enzyme is 4',2',4',6'-  
 CC tetrahydroxychalcone (also termed naringenin-chalcone or chalcone)  
 CC which can under specific conditions spontaneously isomerize into  
 CC naringenin.  
 CC -!- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +  
 CC naringenin-chalcone + 3 CO(2).  
 CC -!- PATHWAY: Part of the biosynthetic pathway for all classes of  
 CC flavonoids, a large class of secondary plant metabolites, many of  
 CC which are brightly colored.  
 CC -!- SIMILARITY: Belongs to the chalcone/stilbene synthases family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; L02905; AAA02827.1; -;  
 CC PIR; S35167; S35167.  
 CC HSP; P30074; IB15.  
 CC InterPro; IPR001099; N-C synthase.  
 CC InterPro; IPR011141; PKS\_III.  
 CC Pfam; PF02797; Chal\_sti\_synt\_C; 1.  
 CC Pfam; PF00195; Chal\_sti\_synt\_N; 1.  
 CC PIRSF; PIRSF000451; PKS\_III; 1.  
 CC ProDom; PD000453; N-C synthase; 1.  
 CC PROSITE; PS00441; CHALCONE\_SYNTH; 1.  
 CC Acyltransferase; Flavonoid biosynthesis; Multigene family;  
 KW Transferrase.  
 FT ACT SITE 164 164 By similarity.  
 FT SEQUENCE 389 AA; 42721 MW; BB461B53B8C63002 CRC64;  
 SQ  
 Query Match 96.4%; Score 1926; DB 1; Length 389;  
 Best Local Similarity 95.9%; Pred. No. 6.3e-131;  
 Matches 373; Conservative 6; Mismatches 10; Indels 0; Gaps 0;  
 QY 1 MVSVEIRKAORAEAGPATILATGTANPANCVEQSTYDPDFYFKITNSEHKTTELKKEKFORMC 60  
 DB 1 MVSVEIRKAORAEAGPATILATGTANPANCVEQSTYDPDFYFKITNSEHKTTELKKEKFORMC 60  
 QY 61 DKSIMKRRYMYLITEILKENPFCVCEYMAPSLDARDQMDVWVVEVPRLGKEAAVKAKEWGQP 120  
 DB 61 DKSIMKRRYMYLITEILKENPFCVCEYMAPSLDARDQMDVWVVEVPRLGKEAAVKAKEWGQP 120  
 QY 121 KSKITHLIVCTTSGVDMPGADYQTLTKLGLRPYKRYMYQOQXCFAGGTVRLAKDLAEN 180  
 DB 121 KSKITHLIFCTTSGVDMPGADYQTLTKLGLRPYKRYMYQOQXCFAGGTVRLAKDLAEN 180  
 QY 181 NKGARVLVVCSEVTATVFRGSDTHLDSLVGOALFGDGAALIVGSDPVPPIEKPIFEMV 240  
 DB 181 NKGARVLVVCSEVTATVFRGSDTHLDSLVGOALFGDGAALIVGSDPVPPIEKPIFEMV 240  
 QY 241 WTAQTIAPDSEGAIDGHLREAGLTFLHLLKDVPGIVSKNITKALVEAFELGIDSYNSIFW 300  
 DB 241 WTAQTIAPDSEGAIDGHLREAGLTFLHLLKDVPGIVSKNIDKALVEAFELGIDSYNSIFW 300  
 QY 301 IAHGPGPAILDQVEQKALKEPKMNAATREVLSEYGNMSSACVLFILDEMRRKSTQNGLKT 360  
 DB 301 IAHGPGPAILDQVEQKALKEPKMNAATREVLSEYGNMSSACVLFILDEMRRKSTQNGLKT 360

```

QY 361 TGGLEWGVLFGRPGGLTITVTLRSVAI 389
DB 361 TGGLEWGVLFGRPGGLTITVTLRSVAI 389

RESULT 12
CHS4_PEA
ID CHS4_PEA STANDARD; PRT; 389 AA.
AC 023882;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Chalcone synthase 4 (EC 2.3.1.74) (Naringenin-chalcone synthase 4).
GN NamesCHS4;
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.
OX NCBI_TaxID=3886;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=cv. Midoriisui;
RA MEDLINE=97374447; PubMed=9230896; DOI=10.1007/s004380050471;
RT Ito M., Ichinose Y., Kato H., Shiraiishi T., Yamada T.;
RT "Molecular evolution and functional relevance of the chalcone synthase
RT genes of pea.";
RL Mol. Gen. Genet. 255:28-37(1997).
CC -!- FUNCTION: The primary product of this enzyme is 4,2',4',6'-
CC tetrahydroxychalcone (also termed naringenin-chalcone or chalcone)
CC which can under specific conditions spontaneously isomerize into
CC naringenin.
CC -!- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +
CC naringenin-chalcone + 3 CO(2).
CC -!- PATHWAY: Part of the biosynthetic pathway for all classes of
CC flavonoids, a large class of secondary plant metabolites, many of
CC which are brightly colored.
CC -!- SIMILARITY: Belongs to the chalcone/stilbene synthases family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; D88260; BAA22042.1; --
CC HSSP; P30074; 1B15.
CC InterPro; IPR001099; N-C_synthase.
CC Pfam; PF02797; Chal_sti_synt_C; 1.
CC Pfam; PF00195; Chal_sti_synt_N; 1.
CC PIRSF; PIRSF000453; PKS_III; 1.
CC ProDom; PD000453; N-C_synthase; 1.
CC PROSITE; PS00441; CHALCONE SYNTH; 1.
KW Acyltransferase; Flavonoid biosynthesis; Multigene family;
FT ACT SITE 164 164 By similarity.
SQ SEQUENCE 389 AA; 42856 MW; 7DCA9D3F266CD125 CRC64;

Query Match 96.4%; Score 1925; DB 1; Length 389;
Best Local Similarity 95.9%; Pred. No. 7.5e-131;
Matches 373; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 1 MUSVSIRKQRAEGPATILAIQTANPANCVEOSTYDPDFYFKITNSEHKTELKEKQFQMC 60
DB 1 MUSVSIRKQRAEGPATILAIQTANPANCVEOSTYDPDFYFKITNSEHKTELKEKQFQMC 60
QY 61 DKSMIKRRYMYLTELKENPNVCYMAPSLDARQDMVVVEVPRLGEAAVKAIKWGGP 120
DB 61 DKSMINRRYMYLTELKENPNVCYMAPSLDARQDMVVVEVPRLGEAAVKAIKWGGP 120
QY 121 KSKITHLIVCTTSGVDMPGADYQLTKLLGLRPPYVKRYMMYQQGCFAGGTVLRKDLAEN 180
DB 121 KSKITHLIVCTTSGVDMPGADYQLTKLLGLRPPYVKRYMMYQQGCFAGGTVLRKDLAEN 180

ID CHS6_TRISU STANDARD; PRT; 389 AA.
AC P51088;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Chalcone synthase 6 (EC 2.3.1.74) (Naringenin-chalcone synthase 6).
GN NamesCHS6;
OS Trifolium subterraneum (Subterranean clover).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Trifolium.
OX NCBI_TaxID=3900;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=cv. Karridale; TISSUE=Leaf, and Stem;
RA MEDLINE=95232194; PubMed=7716240; DOI=10.1104/pp.107.3.1035;
RT Howles P.A., Arioli T., Weinman J.J.;
RT "Nucleotide sequence of additional members of the gene family encoding
RT chalcone synthase in Trifolium subterraneum.";
RL Plant Physiol. 107:1035-1036(1995).
CC -!- FUNCTION: The primary product of this enzyme is 4,2',4',6'-
CC tetrahydroxychalcone (also termed naringenin-chalcone or chalcone)
CC which can under specific conditions spontaneously isomerize into
CC naringenin.
CC -!- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +
CC naringenin-chalcone + 3 CO(2).
CC -!- PATHWAY: Part of the biosynthetic pathway for all classes of
CC flavonoids, a large class of secondary plant metabolites, many of
CC which are brightly colored.
CC -!- INDUCTION: By wounding and Rhizobium infection.
CC -!- SIMILARITY: Belongs to the chalcone/stilbene synthases family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; M91195; AAA67701.1; --
CC HSSP; P30074; 1B15.
CC InterPro; IPR001099; N-C_synthase.
CC InterPro; IPR01141; PKS_III.
CC Pfam; PF02797; Chal_sti_synt_C; 1.
CC Pfam; PF00195; Chal_sti_synt_N; 1.
CC PIRSF; PIRSF000451; PKS_III; 1.
CC ProDom; PD000453; N-C_synthase; 1.
CC PROSITE; PS00441; CHALCONE SYNTH; 1.
KW Acyltransferase; Flavonoid biosynthesis; Multigene family;
KW Transferase.

```

FT ACT\_SITE 164 164 By similarity.  
SQ SEQUENCE 389 AA; 42672 MW; A67A9155A89E68E4 CRC64;  
Query Match 96.1%; Score 1920; DB 1; Length 389;  
Best Local Similarity 95.9%; Pred. No. 1.7e-130;  
Matches 373; Conservative 7; Mismatches 9; Indels 0; Gaps 0;  
QY 1 MVSSEIRKQAEAGPATILAIAGTANPANCVEQSTYDPFYFKITNSEHKTTELKEKFORMC 60  
Db 1 MVSVAEIRKQAEAGPATILAIAGTANPANKVEQATYDPFYFKITNSEHKTTELKEKFORMC 60  
QY 61 DKSMIKRMYLTYTEILKENPNVCEYMAPSLDARDQMDVVPRLGKEAAVKAKEWGP 120  
Db 61 DKSMIKRMYLTYTEILKENPSLCEYMAPSLDARDQMDVVPRLGKEAAVKAKEWGP 120  
QY 121 KSKITHLIVCTTSGVDMPCADYQTLKGLRPVVKYMYQGXFGAGTVLRKADLAEN 180  
Db 121 KSKITHLIFCTTSGVDMPCADYQTLKGLRPVVKYMYQGXFGAGTVLRKADLAEN 180  
QY 181 NKGARVLVVCSEVTAATFRGPDSTHLDLSLVGQALFGDGAALIVGSDPVPPIEFMV 240  
Db 181 NKGARVLVVCSEVTAATFRGPDSTHLDLSLVGQALFGDGAALIVGSDPVPPIEFMV 240  
QY 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFPLGSDYNSIFW 300  
Db 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNIDKALVEAFQPLNIDYNSIFW 300  
QY 301 IAHPGGPAILDQVEOKLALPEKKNATREVLSEYGNMSSACVLFILDEMREKSTONGLKT 360  
Db 301 IAHPGGPAILDQVEOKLALPEKKNATREVLSEYGNMSSACVLFILDEMREKSTONGLKT 360  
QY 361 TGEGLWGVLFPGFGPLTIETVLRSAI 389  
Db 361 TGEGLWGVLFPGFGPLTIETVLRSAI 389  
RESULT 14  
CHS6\_PEA  
ID CHS6\_PEA STANDARD; PRT; 389 AA.  
AC Q01288;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Chalcone synthase 6 (EC 2.3.1.74) (Naringenin-chalcone synthase 6).  
GN Names=CHS6;  
OS Pisum sativum (Garden pea).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosid1; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.  
OX NCBI\_TaxID=3888;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Midori; TISSUE=epicotyl;  
RX MEDLINE=92256801; PubMed=1581561;  
RA Ichinose Y., Kawamata S., Yamada T., An C., Kajiura T., Shiraishi T.,  
RA Oku H.;  
RT "Molecular cloning of chalcone synthase cDNAs from Pisum sativum.";  
RL Plant Mol. Biol. 18:1009-1012(1992).  
CC -!- FUNCTION: The primary product of this enzyme is 4,2',4',6'-  
CC tetrahydrochalcone (also termed naringenin-chalcone or chalcone)  
CC which can under specific conditions spontaneously isomerize into  
CC naringenin.  
CC -!- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +  
CC naringenin-chalcone + 3 CO(2).  
CC -!- PATHWAY: Part of the biosynthetic pathway for all classes of  
CC flavonoids, a large class of secondary plant metabolites, many of  
CC which are brightly colored.  
CC -!- SIMILARITY: Belongs to the chalcone/stilbene synthases family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X63335; CAA44935.1; -.  
DR PIR; S20933; S20933.  
DR HSSP; P30074; 1CGZ.  
DR InterPro; IPR001099; N-C synthase.  
DR InterPro; IPR011141; PKS\_III.  
DR Pfam; PF02797; Chal\_sti\_synt\_C; 1.  
DR Pfam; PF00195; Chal\_sti\_synt\_N; 1.  
DR PIRSF; PIRSF000451; PKS\_III; 1.  
DR ProDom; PD000453; N-C synthase; 1.  
DR PROSITE; PS00441; CHALCONE SYNTH; 1.  
KW Acyltransferase; Flavonoid biosynthesis; Multigene family;  
KW Transferase.  
FT ACT\_SITE 164 164 By similarity.  
SQ SEQUENCE 389 AA; 42761 MW; 73FC26AFDF82EEDD CRC64;  
Query Match 95.9%; Score 1915; DB 1; Length 389;  
Best Local Similarity 95.4%; Pred. No. 4e-130;  
Matches 371; Conservative 6; Mismatches 12; Indels 0; Gaps 0;  
QY 1 MVSSEIRKQAEAGPATILAIAGTANPANCVEQSTYDPFYFKITNSEHKTTELKEKFORMC 60  
Db 1 MVSSEIRKQAEAGPATILAIAGTANPANCVEQSTYDPFYFKITNSEHKTTELKEKFORMC 60  
QY 61 DKSMIKRMYLTYTEILKENPNVCEYMAPSLDARDQMDVVPRLGKEAAVKAKEWGP 120  
Db 61 DKSMIKRMYLTYTEILKENPSLCEYMAPSLDARDQMDVVPRLGKEAAVKAKEWGP 120  
QY 121 KSKITHLIVCTTSGVDMPCADYQTLKGLRPVVKYMYQGXFGAGTVLRKADLAEN 180  
Db 121 KSKITHLIFCTTSGVDMPCADYQTLKGLRPVVKYMYQGXFGAGTVLRKADLAEN 180  
QY 181 NKGARVLVVCSEVTAATFRGPDSTHLDLSLVGQALFGDGAALIVGSDPVPPIEFMV 240  
Db 181 NKGARVLVVCSEVTAATFRGPDSTHLDLSLVGQALFGDGAALIVGSDPVPPIEFMV 240  
QY 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFPLGSDYNSIFW 300  
Db 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNINKALVEAFQPLNIDYNSIFW 300  
QY 301 IAHPGGPAILDQVEOKLALPEKKNATREVLSEYGNMSSACVLFILDEMREKSTONGLKT 360  
Db 301 IAHPGGPAILDQVEOKLALPEKKNATREVLSEYGNMSSACVLFILDEMREKSTONGLKT 360  
QY 361 TGEGLWGVLFPGFGPLTIETVLRSAI 389  
Db 361 TGEGLWGVLFPGFGPLTIETVLRSAI 389  
RESULT 15  
CHS5\_TRISU  
ID CHS5\_TRISU STANDARD; PRT; 389 AA.  
AC P51087;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Chalcone synthase 5 (EC 2.3.1.74) (Naringenin-chalcone synthase 5).  
GN Name=CHS5;  
OS Trifolium subterraneum (Subterranean clover).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosid1; Fabales; Fabaceae; Papilionoideae; Trifolieae; Trifolium.  
OX NCBI\_TaxID=3900;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Karri-dale; TISSUE=Leaf, and Stem;  
RX MEDLINE=95232194; PubMed=7716240; DOI=10.1104/pp.107.3.1035;  
RA Howles P.A., Arioli T., Weinman J.J.;  
RT "Nucleotide sequence of additional members of the gene family encoding  
RT chalcone synthase in Trifolium subterraneum.";

Plant Physiol. 107:1035-1036(1995)).  
-!- FUNCTION: The primary product of this enzyme is 4,2',4',6'-  
tetrahydrochalcone (also termed naringenin-chalcone or Chalcone)  
which can under specific conditions spontaneously isomerize into  
naringenin.  
-!- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +  
naringenin-chalcone + 3 CO(2).  
-!- PATHWAY: Part of the biosynthetic pathway for all classes of  
flavonoids, a large class of secondary plant metabolites, many of  
which are brightly colored.  
-!- INDUCTION: By wounding and Rhizobium infection.  
-!- SIMILARITY: Belongs to the chalcone/stilbene synthases family.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; L24517; AAU73939.1; -.  
DR HSP; P30074; 1B15  
DR InterPro; IPR01099; N-C\_synthase.  
DR InterPro; IPR01141; PKS\_III.  
DR Pfam; PF02797; Chal\_sti\_synt\_C; 1.  
DR Pfam; PF00195; Chal\_sti\_synt\_N; 1.  
DR PIRSF; PIRSF00451; PKS\_III; 1.  
DR PRODOM; PD000453; N-C\_synthase; 1.  
DR PROSITE; PS00441; CHALCONE\_SYNTH; 1.  
KW Acyltransferase; Flavonoid biosynthesis; Multigene family;  
FT ACT SITE 164 164 BY similarity.  
SQ SEQUENCE 389 AA; 42656 MW; EEAC4AA9717D4CA9 CRC64;  
  
Query Match 95.8%; Score 1914; DB 1; Length 389;  
Best Local Similarity 95.6%; Pred. No. 4.7e-130;  
Matches 372; Conservative 7; Mismatches 10; Indels 0; Gaps 0;  
  
QY 1 MVSVEIRKAQRAEGPATILAICTANPANCVEQSTYDPDFYFKITNSEHKTKEKFORMC 60  
DB 1 MVSVAIRKAQRAEGPATILAICTANPANKVEQATYDPDFYFKITNSEHKTKEKFORMC 60  
  
QY 61 DKSMIKRRYMYLTPEEILKENPNVCEYMAPSLDARQDMVVVEVPRLGKEAAVKAKEWGQP 120  
DB 61 DKSMIKRYMYLTPEEILKENPSVCEYMAPSLDARQDMVVVEVPRLGKEAAVKAKEWGQP 120  
  
QY 121 KSKITHLIVCTTSGVDMPGADYQLTKLGLRPYVKRYMMYQQXFPAGGTVLRKDLAEN 180  
DB 121 KSKITHLIFCTTSGVDMPGADYQLTKLGLRPYVKRYMMYQQCFAGGTVLRKDLAEN 180  
  
QY 181 NKGARVLVWCSEVTATFRGPDTHLDSLQALFGDGAALIVGSDVPPEIEKPIFEMV 240  
DB 181 NKGARVLVWCSEVTATFRGPDTHLDSLQALFGDGAALIVGSDVPPEIEKPIFEMV 240  
  
QY 241 WTAQTITAPDSEGAIDGHLREAGLTFFHLKDVPCIVSKNITKALVEAFPLGIDSYNSIFW 300  
DB 241 WTAQTITAPDSEGAIDGHLREAGLTFFHLKDVPCIVSKNIDKALVEAFPLNSIDYNSIFW 300  
  
QY 301 IAHPGGPAILDQVEQKALKPEKONATREVLSEYGNMSSACVLFILDEMRRKKSQTGLKLT 360  
DB 301 IAHPGGPAILDQVEQKALKPEKOKATRDVLSYGNMSSACVLFILDEMRRKKSQAQGLKLT 360  
  
QY 361 TGEGLWGVLPFGPGGLTITETVLRVAI 389  
DB 361 TGEGLDWGVLPFGPGGLTITETVLRVAI 389

Search completed: April 22, 2005, 21:44:16  
Job time : 174 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2005, 21:35:47 ; Search time 42 Seconds  
(without alignments)  
691.393 Million cell updates/sec

Title: US-10-031-918A-1  
Perfect score: 1997  
Sequence: 1 MVSVSIRKAQRAEGPATIL.....LFGFGELTETVLRSAI 389

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PTUS COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1508.5	75.5	394	US-09-252-816A-1	Sequence 1, Appli
2	1508.5	75.5	394	US-09-837-654-1	Sequence 1, Appli
3	477	23.9	151	US-08-321-358-2	Sequence 2, Appli
4	314	15.7	349	US-08-994-035C-5	Sequence 5, Appli
5	314	15.7	349	US-09-395-861-5	Sequence 5, Appli
6	310	15.5	349	US-08-494-907-12	Sequence 12, Appl
7	310	15.5	349	PCT-US96-10986-12	Sequence 12, Appl
8	304	15.2	360	US-09-902-540-13778	Sequence 13778, A
9	183	9.2	506	US-09-877-476-20	Sequence 20, Appl
10	181	9.1	506	US-09-877-476-16	Sequence 16, Appl
11	181	9.1	506	US-09-877-476-22	Sequence 22, Appl
12	180	9.0	506	US-09-877-476-40	Sequence 40, Appl
13	179.5	9.0	521	US-08-657-749D-6	Sequence 6, Appli
14	179	9.0	506	US-09-877-476-12	Sequence 12, Appl
15	178	8.9	505	US-09-877-476-6	Sequence 6, Appli
16	177	8.9	506	US-09-877-476-8	Sequence 8, Appli
17	177	8.9	506	US-09-877-476-14	Sequence 14, Appl
18	177	8.9	506	US-09-877-476-38	Sequence 38, Appl
19	177	8.9	506	US-09-877-476-42	Sequence 42, Appl
20	176	8.8	475	US-08-657-749D-11	Sequence 11, Appl
21	176	8.8	475	US-08-657-749D-13	Sequence 13, Appl
22	176	8.8	505	US-09-877-476-13	Sequence 18, Appl
23	176	8.8	506	US-09-877-476-26	Sequence 26, Appl
24	176	8.8	506	US-09-877-476-34	Sequence 34, Appl
25	173	8.7	506	US-09-877-476-24	Sequence 24, Appl
26	173	8.7	506	US-09-877-476-32	Sequence 32, Appl
27	172	8.6	505	US-09-877-476-10	Sequence 10, Appl

28	172	8.6	506	4	US-09-877-476-4	Sequence 4, Appli
29	171	8.6	506	3	US-08-888-998-2	Sequence 2, Appli
30	171	8.6	506	3	US-09-362-633-2	Sequence 2, Appli
31	171	8.6	506	4	US-09-877-476-2	Sequence 2, Appli
32	171	8.6	506	4	US-09-877-476-28	Sequence 28, Appli
33	170	8.5	506	4	US-09-877-476-30	Sequence 30, Appli
34	169	8.5	506	4	US-09-877-476-36	Sequence 36, Appli
35	167	8.4	333	3	US-09-134-001C-3946	Sequence 3946, Ap
36	166	8.3	493	3	US-08-868-373-4	Sequence 4, Appli
37	165	8.3	524	4	US-08-657-749D-4	Sequence 4, Appli
38	164	8.2	505	4	US-08-657-749D-25	Sequence 25, Appli
39	162.5	8.1	497	3	US-09-058-947A-4	Sequence 4, Appli
40	162.5	8.1	500	3	US-08-868-373-12	Sequence 12, Appli
41	156.5	7.8	516	3	US-08-868-373-14	Sequence 14, Appli
42	156	7.8	498	4	US-08-657-749D-23	Sequence 23, Appli
43	155	7.8	313	4	US-08-970-647-2	Sequence 2, Appli
44	154.5	7.7	504	3	US-08-868-373-6	Sequence 6, Appli
45	151	7.6	313	4	US-10-138-701-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1  
US-09-252-816A-1  
; Sequence 1, Application US/09252816A  
; Patent No. 6265633  
; GENERAL INFORMATION:  
; APPLICANT: OKADA, Yukio  
; APPLICANT: ITO, Kazutoshi  
; TITLE OF INVENTION: ISOLATED AND PURIFIED NUCLEIC ACIDS COMPRISING A GENE  
; TITLE OF INVENTION: AND A REGULATORY REGION FOR THE GENE EXPRESSION OF THE  
; TITLE OF INVENTION: SAME  
; FILE REFERENCE: 1959-0008-0  
; CURRENT APPLICATION NUMBER: US/09/252,816A  
; CURRENT FILING DATE: 1999-02-19  
; PRIOR APPLICATION NUMBER: JP HEI 10-37266  
; PRIOR FILING DATE: 1998-02-19  
; PRIOR APPLICATION NUMBER: JP HEI 10-174235  
; PRIOR FILING DATE: 1998-06-22  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 394  
; TYPE: PRT  
; ORGANISM: Humulus lupulus  
US-09-252-816A-1

Query Match	75.5%	Score 1508.5;	DB 3;	Length 394;
Best Local Similarity	72.1%	Pred. No. 5.2e-153;		
Matches 279;	Conservative 59;	Mismatches 48;	Indels 1;	Gaps 1;
QY	2	VSVSEIRKAQRAEGPATILAIAGTANPANCVEQSTYDPYFKITNSHKTELKKEKFORMCD 61		
Db	4	VTVQIRKAQRAEGPATILAIAGTAVPANCFCNOADFPDYFRTVKSEHMTDLKKKFORMCE 63		
QY	62	KSMIKERYMLTEETILKENPNVCEYMAPSLDARQDMVVVEPRLGEEAAVKAIKEGQPK 121		
Db	64	KSTIKRYLHLTEEHUKQNPILCEYNAPSLNRQDMVLVVEVPKLGKEAINALKEGQPK 123		
QY	122	SKITHLIVCTSGVDMFGADYQTLKLLGRPPYKRYMMYQQGXFAGTGLRLAKDAENN 181		
Db	124	SKITHLIFCTGSSIDMPGADYQCAKLLGLRPSVKRYMLYQLGCGYAGKVLRIAKDAENN 183		
QY	182	KGARVLVCSVTAFTFRGFSPTHLDLSVQALFGGAAALIVGSPVPEI-EKPIPEMV 240		
Db	184	KGARVLVCSSEITACIFRGFSEKHLDCVLGQSLFGDGASSVIVGADPDASVGERPFELV 243		
QY	241	WTAQTTAPSEGAIDGHLREAGLTTHLLKDVGVISKNITKALVEAFEPGLGSDYNSIFW 300		
Db	244	SAAQTTLPNSDGAIGHVTEAGLTTHLLRDVPLGLISQNIKSLIEAFTPIGINDMNNIFW 303		
QY	301	IAHPGGPAILDQVEQKALKPEKMNATREVLSEYGNSSACVLFILDEMRKIKSTQNGLKT 360		

Db	Qy	Db
304	IAHPGPAALDIEAKLELKKEKKASREMLSEYGNNSCASVFIIVDEMRQSSKEGKST	363
361	TGEGLEWGVLCFGFGPLTIETVVLRSV	387
364	TGEGLEWGVLCFGFGPLTIETVVLHSV	390

RESULT 2  
US-09-837-654-1  
; Sequence 1, Application US/09837654  
; Patent No. 6639127  
; GENERAL INFORMATION:  
; APPLICANT: OKADA, Yukio  
; APPLICANT: ITO, Kazutoshi  
; TITLE OF INVENTION: ISOLATED AND PURIFIED NUCLEIC ACIDS COMPRISING A GENE  
; TITLE OF INVENTION: AND A REGULATORY REGION FOR THE GENE EXPRESSION OF THE  
; TITLE OF INVENTION: SAME  
; FILE REFERENCE: 1959-0008-0  
; CURRENT APPLICATION NUMBER: US/09/837,654  
; CURRENT FILING DATE: 2001-04-19  
; EARLIER APPLICATION NUMBER: 09/252,816  
; EARLIER FILING DATE: 1999-02-19  
; EARLIER APPLICATION NUMBER: JP HEI 10-174235  
; EARLIER FILING DATE: 1998-06-22  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 394  
; TYPE: PRT  
; ORGANISM: Humulus lupulus  
US-09-837-654-1

Query Match	75.5%;	Score 1508.5;	DB 4;	Length 394;
Best Local Similarity	72.1%;	Pred. No. 5.2e-153;		
Matches 279;	Conservative	59;	Mismatches 48;	Indels 1;
				Gaps 1;

Qy	2	VSVSEIRKAQRAEGPATIIAIGTANPANCVEQSTYPPDYFKIUNSEHKTELKCKFORMCD	61
Db	4	VTVQIRKAQRAEGPATIIAIGTAVPANCQADFPDYFRVTKSEHMTDLKKFFORMCE	63
Qy	62	KSMIKRMYMLTEEBILKENPNVCYMAPSLDARODMVVVEPRLKGEEAAVKAIKENGQPK	121
Db	64	KSTIKRYLHUTEELKGNPHLCEYNAPSLNTRQDMLVVEPKLKGEEAANAIKENGQPK	123
Qy	122	SKITHLIVCTTSVDMPGADYQLTKLLGLREYVVKRYMMYQCGXFGGTVLRLAKDLAENN	181
Db	124	SKITHLIFCTGSSIDMPGADYQCAKLLGLRFSVKRVMLYQLGCVAGGKVLRIAKDLAENN	183
Qy	182	KGARVLVGCSEVATVTFRGPSDTHLSVLQALFGDGAALIVGSDVPVPI-BKPIPEMV	240
Db	184	KGARVLVGCSEITACIFRGPSEKHLDCILVGQSLFGDGASSVVCADPDASVGERPIFELV	243
Qy	241	WTAQTIAPDSEGAIDGHLREAGTLFHLKDPVGIKSNITKALVEAPELIGISDYNISFW	300
Db	244	SAAQTIPLNSDGAIAGHVTEAGTLFHLRDPVGLISQNIIEKSLTEAFTPIGINDWNISFW	303
Qy	301	IAHGGPAILDVOEQKLALKEPKKNATREVLSEYGNMSSACVLFIIDEMRKKSQTONGLKT	360
Db	304	IAHGGPAILDEIEAKLELKEKKKASREMLSEYGNMSSACSVFVIDEMRKQSSKSGKST	363
Qy	361	TGEGLEWGVLFQFGPLTIETIVLRSV	387
Db	364	TGEGLEWGVLFQFGPLVETIVLHVSV	390

RESULT 3  
US-08-321-358-2  
; Sequence 2, Application US/08321358  
; Patent No. 5589620  
; GENERAL INFORMATION:  
; APPLICANT: Helmut Kindl; Rudiger Hain; Hans-Jorg Reif  
; TITLE OF INVENTION: B1ENZYL SYNTHASE GENES

NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SPRUNG HORN KRAMER & WOODS  
STREET: 660 White Plains Road  
CITY: Tarrytown  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10591-5144

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB

COMPUTER: HP VECTRA

SOFTWARE: Wordperfect 5.1

APPLICATION NUMBER: US/08/321,358  
FILING DATE: 11-OCT-1994

CLASSIFICATION: 435  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: P 43 34 /91.6 (German)  
FILING DATE: 13-OCT-1993

ATTORNEY/AGENT INFORMATION:  
NAME: Kurt G. Briscoe

REFERENCE/DOCKET NUMBER: Bayer 9137-KGB

TELEPHONE: (914) 332-1700  
TELETYPE: (914) 332-1844

TELEX:  
INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:  
LENGTH: 151 Amino acids

**LIFE: AMINO ACID**  
**TOPOLOGY: Linear**

3-321-358-2

Very Match 23.9%; Score 477; DB  
at local similarity: prod No 6047

Matches 86; Conservative 30; Mismatches

1 MVSVSEIRKAQRAEGPATILLAGTANPANCVEQS

1 MFSDS1NAAFKADGEAS1EALAGNINEUNITES

[illegible]

1 2 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 10

121 KSPITHI.IBC.TTSCMNI.BGADYOL.TOT 147

3-994-035C-5

Agent No. 6277625

APPLICANT: Huang, Zhengyu

APPLICANT: Mavrodi, Dmitri V  
APPLICANT: Paalimakers, Jos M

APPLICANT: Weller, David M  
APPLICANT: Cook, R James

TITLE OF INVENTION:	TRANSGENIC STRAINS FOR
TITLE OF INVENTION:	PLANT ROOT DISEASES

CORRESPONDENCE ADDRESS:

STREET: 800 Buchanan St  
CITY: Atlanta

STATE: CA



```

; LENGTH: 349 amino acids
; TYPE: amino acid
;
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-395-861-5

```

Query Match	15.7%	Score 314;	DB 4;	Length 349;
Best Local Similarity	27.6%;	Pred. No. 7.4e-25;		
Matches 103;	Conservative 59;	Mismatches 169;	Indels 42;	Gaps 9;
QY	30	CVEQSYPPDYFKITNS---EHTKELKEF-----QRMCDSKMIKRYMYLVEILKEN	80	
DB	5	CRPSLLFP--HYKIQOQIMIDHLEQJLDDHPRMALAKRMTQNTQVNERIVLVPIDELAVH	62	
QY	81	PNVCEYMAPSLDARQDMVVVEVPLGKEAAVKAIKEWGQPKSKITHLIVCTTSQVDMPGA	140	
DB	63	TGETH-----RSIVYERARMSSIAQRAQIENAGLTTDDIRMVAVVTSCTGFMPPSL	114	
QY	141	DYQLTKLLGLRPPVYKRYMYQQGXFGAGTVLRLAKDLAENKNGARVLVVCSEVTAFTFRG	200	
DB	115	TAHLINDLGLRTSTVQLPTAQLSCVAGAAAINRANDFASLSPDNHALIVLSLEPSSLCVQ-	173	
QY	201	PSDTHLDLSVGQALFGDGAALIVGSDPVEIEKPIEFEMVWTQTATPADSEGAIDGHLRE	260	
DB	174	PDOTKLHAFITSAIFGDAVSACVMRADD----KAPGFKIAKTGSYFLPDSHYHIKYQVKD	229	
QY	261	AGLTFHLKDVPGIVSKNITHKALVEAFELIGISDYN----SIFWIAHPCGPAILDDVEQ	315	
DB	230	SGFHFTLDKAV-----MNSIKDVAPMWEELNFTFNQHCQNDFPIFHTGGRKLLDEVL	284	
QY	316	KLALKEPKGNATREVLSEYGNMSSACVLFILDEMRKKSTQNGLKTTEGGEWGLVFGFGP	375	
DB	285	QLDLEPGRVAQSRDSLSEAGNIASVVVFDVLKQFDSGPGANGAT-----GMLAAGFP	336	
QY	376	GLTIEVTVVLRSA	388	
DB	337	GFTAEMAVGKWVA	349	

QY 376 GLTIETVWLRVA 388  
| | | | : | |  
| | | | : | |  
Db 337 GFTAEMAVGKWA 349

RESULT 6  
US-08-494-907-12  
; Sequence 12, Application US/08494907  
; Patent No. 595298  
; GENERAL INFORMATION:  
; APPLICANT: Thomasow, Linda S  
; APPLICANT: Bangera, Mahalaxmi  
; APPLICANT: Weller, David M  
; APPLICANT: Cook, R. James

TITLE OF INVENTION: Sequences for Production of  
TITLE OF INVENTION: 2,4-Diacetylphloroglucinol and Methods  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Margaret A. Connor, USDA-ARS  
STREET: 800 Buchanan Street  
CITY: Albany  
STATE: CA  
COUNTRY: USA  
ZIP: 94710

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/494,907  
FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Connor, Margaret A

REGISTRATION NUMBER: 30043

REFERENCE/DOCKET NUMBER: 0009.95

TELECOMMUNICATION INFORMATION:

TELEPHONE: (510) 559-6067

TELEFAX: (510) 559-5777

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 349 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

PCT-US96-10986-12

US-08-494-907-12

Query Match 15.5%; Score 310; DB 2; Length 349;

Best Local Similarity 27.3%; Pred. No. 2e-24;

Matches 102; Conservative 59; Mismatches 170; Indels 42; Gaps 9;

QY 30 CVEQSTYDPFYKITS---EHKTELKEF-----QRMCDKSMIKRYMYLTELKEN 80

Db 5 CKPSLLFP--HYKITQQQMHLEQLHDDHPRMALAKMIQNTQVNERVLPIDELAVH 62

QY 81 PNVCYMAPSLDARDQMVVVEPRLGKEAAVKAKEWGQPKSKITHLIVCTTSGVDMPGA 140

Db 63 TGFTH-----RSIVYERARRMSSIAARQAIEAENAGLTDDIRWAVTSGTFMPSL 114

QY 141 DYQTLKLGRLPVVKRYMYQGXFGAGTVLRKADLAENKRGARVLVVCSEVTAFTFG 200

Db 115 TAHLINDLGLRTSTVQLPIAQLGCVAGAAAINRANDFGSLSPDNHALIVSLFSSLCYQ- 173

QY 201 PSDTHLDSLVGOALFGDGAALIVGSDVPVEIEKPIFEMVWTAQTAPDSEGAIDHLE 260

Db 174 PQDTKLHAFISAALFGDAVSACVWRADD---KAPGFKIAKTGSYFLPSEHYIKYDVKD 229

QY 261 AGLTFLHLKDVPGVIGSKNITKALVEAFEPGLGISDYN-----SIFWIAHPGGPAILDQVEQ 315

Db 230 SGHFETLDRAV-----MNSIKDVAAPMMEELNFTFNQHCANDFFIHTGGRKILDELVL 284

QY 316 KLAKEPKMNAITREVISEYGNMSSACVLFILDEMRKKSTQNGLKTGEGLEWGLFGFQ 375

Db 285 QLDLEPGVAQSRDLSSEAGNIASVVVFDVLRQFDSGPANGAT-----GMLAAGFP 336

QY 376 GLTIETVILRSVA 388

Db 337 GFTAEMAVGKWA 349

RESULT 7

PCT-US96-10986-12

Sequence 12, Application PC/TUS9610986

GENERAL INFORMATION:

TITLE OF INVENTION: Sequences for Production of

TITLE OF INVENTION: 2,4-Diacetylphloroglucinol and Methods

NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Stephan A. Pendorf, DOMINIK & STEIN  
STREET: 600 N. West Shore Boulevard, Suite 1000  
CITY: Tampa  
STATE: FL  
COUNTRY: USA  
ZIP: 33609

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/10986  
FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Pendorf, Stephan A.

REGISTRATION NUMBER: 32665

REFERENCE/DOCKET NUMBER: A700.320

TELECOMMUNICATION INFORMATION:

TELEPHONE: (813) 289-2966

TELEFAX: (813) 289-2967

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 349 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

PCT-US96-10986-12

Query Match 15.5%; Score 310; DB 5; Length 349;

Best Local Similarity 27.3%; Pred. No. 2e-24;

Matches 102; Conservative 59; Mismatches 170; Indels 42; Gaps 9;

QY 30 CVEQSTYDPFYKITS---EHKTELKEF-----QRMCDKSMIKRYMYLTELKEN 80

Db 5 CKPSLLFP--HYKITQQQMHLEQLHDDHPRMALAKMIQNTQVNERVLPIDELAVH 62

QY 81 PNVCYMAPSLDARDQMVVVEPRLGKEAAVKAKEWGQPKSKITHLIVCTTSGVDMPGA 140

Db 63 TGFTH-----RSIVYERARRMSSIAARQAIEAENAGLTDDIRWAVTSGTFMPSL 114

QY 141 DYQTLKLGRLPVVKRYMYQGXFGAGTVLRKADLAENKRGARVLVVCSEVTAFTFG 200

Db 115 TAHLINDLGLRTSTVQLPIAQLGCVAGAAAINRANDFGSLSPDNHALIVSLFSSLCYQ- 173

QY 201 PSDTHLDSLVGOALFGDGAALIVGSDVPVEIEKPIFEMVWTAQTAPDSEGAIDHLE 260

Db 174 PQDTKLHAFISAALFGDAVSACVWRADD---KAPGFKIAKTGSYFLPSEHYIKYDVKD 229

QY 261 AGLTFLHLKDVPGVIGSKNITKALVEAFEPGLGISDYN-----SIFWIAHPGGPAILDQVEQ 315

Db 230 SGHFETLDRAV-----MNSIKDVAAPMMEELNFTFNQHCANDFFIHTGGRKILDELVL 284

QY 316 KLAKEPKMNAITREVISEYGNMSSACVLFILDEMRKKSTQNGLKTGEGLEWGLFGFQ 375

Db 285 QLDLEPGVAQSRDLSSEAGNIASVVVFDVLRQFDSGPANGAT-----GMLAAGFP 336

QY 376 GLTIETVILRSVA 388

Db 337 GFTAEMAVGKWA 349

RESULT 8

US-09-902-540-13778

Sequence 13778, Application US/09902540

Patent No. 6833447

GENERAL INFORMATION:

APPLICANT: Goldman, Barry S.

APPLICANT: Hinkle, Gregory J.

APPLICANT: Slater, Steven C.

[illegible]

```

RESULT 10
US-09-877-476-16
; Sequence 16, Application US/09877476
; Patent NO. 6713664
; GENERAL INFORMATION:
; APPLICANT: Jaworski, Jan G.
; APPLICANT: Blacklock, Brenda J.
; TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA
; TITLE OF INVENTION: SYNTHASE POLYPEPTIDES
; FILE REFERENCE: 07148-108001
; CURRENT APPLICATION NUMBER: US/09/877,476
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,326
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: 5' 114 amino acids from A. thaliana FAE1 (SEQ ID
; OTHER INFORMATION: NO:2) and 3' 392 amino acids from B. napus
; OTHER INFORMATION: elongase KCS (SEQ ID NO:4) having mutation at
; OTHER INFORMATION: residue 307; designated At114 G307D; hypothetical
; US-09-877-476-16

```

```

; PRIOR APPLICATION NUMBER: US 60/210,326
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: 5' 114 amino acids from A. thaliana FAE1 (SEQ ID
; OTHER INFORMATION: NO:2) and 3' 392 amino acids from B. napus
; OTHER INFORMATION: elongase KGS (SEQ ID NO:4) having mutation at
; OTHER INFORMATION: residue 307; designated At114 G307D; hypothetical
US-09-877-476-16

Query Match          9 1%; Score 181; DB 4; Length 506;
Best Local Similarity 22.9%; Pred. NO. 2.6e-10;
Matches 95; Conservative 63; Mismatches 159; Indels 98; Gaps 18

Qy 26 NPANCVBQSTV-----PDVFVKI-----TNSEHKTELKEKFORMCDKS----- 63
Db 75 NPVYLVDSYCYLPPPHLKVSUSKWDIFYQIRKADTSRNGT-----CDDSSWLDF 125
Qy 64 --MIKRR-----YMYLFEELKENPNVCEYMAPSLDARQDMVVEVPLKGEAAVKAIKE 116
Db 126 LRKIOERSGLGDETHGPEGLIQVPPR--KTFPAAREETEQVYIIIGALENLFNNTV----- 178

```

QY 117 WGPKSKITHLIIVCTTSGVDMPCADYQLTKLLGLRPVVKRYMMYQQXFGAGTVLR LAKD 176  
Db 179 --NPKD-IGILVNSSWFNFTPSLSAMVNTFKLRNVRNFRNLGGMGCSAGVIAIDLAKD 235  
QY 177 LAENNKGARVLVVCSEVTAATFRGSDTHLDSLVGOALFGDGAALIVGSDPVPETKEPI 236  
Db 236 LLVHKNTVALVSTE--NITYNIYAGDNRMVSNCLFRVGGAAILLNKP-GDRRRSK 292  
QY 237 FEMVWTAQTIAPDSEGAIDGHLR-----EAG-LTFHLLKDVPGI-----VSKNIT- 280  
Db 293 YELVHTVRT-----HTGADDKSFRVCQQGDDENGKIGVSLSKDITDVAGRTVKKNIA TLGP 348  
QY 281 -----KALVEAFEPGLGSDYNSIF--WIAHPGCPAILDQVEQKLAUK 320  
Db 349 LILPLSEKLLFFVTMGKLLFKDKIKHYVVPDFKLAIDHFCIHAGGRAVIDVLEKNLALA 408  
QY 321 PERONATREVLSEYGNMSSACVLFILDEMRKKS-TONGLKTTTGEGLWGLFGFG 374  
Db 409 PIDVEASRSTLHRFGNTSSSIWYELAYIEAKGRMKGNKV-----WQIALGSG 457

## RESULT 11

US-09-877-476-22

; Sequence 22, Application US/09877476

; Patent No. 6713664

; GENERAL INFORMATION:

; APPLICANT: Jaworski, Jan G.

; APPLICANT: Blacklock, Brenda J.

; TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA

; TITLE OF INVENTION: SYNTHASE POLYPEPTIDES

; FILE REFERENCE: 07148-108001

; CURRENT APPLICATION NUMBER: US/09/877,476

; CURRENT FILING DATE: 2001-06-08

; PRIOR APPLICATION NUMBER: US 60/210,326

; PRIOR FILING DATE: 2000-06-08

; NUMBER OF SEQ ID NOS: 56

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 22

; LENGTH: 506

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: 5' 114 amino acids from A. thaliana FAEL (SEQ ID

; OTHER INFORMATION: No:2) and 3' 392 amino acids from B. napus

; OTHER INFORMATION: elongase KCS (SEQ ID NO:4) having mutations at

; OTHER INFORMATION: positions 92 and 307; designated At114 K92R G307D;

; OTHER INFORMATION: hypothetical

US-09-877-476-22

Query Match 9.1%; Score 181; DB 4; Length 506;

Best Local Similarity 22.9%; Pred. No. 2.6e-10; Indels 98; Gaps 18;

Matches 95; Conservative 63; Mismatches 159; Indels 98; Gaps 18;

QY 26 NPANCVEQSTY-----PDFYFKI-----TNSEHTELKEKFORMCDKS----- 63

Db 75 NPVYLVDYSCYLPPLHRLVSVSKVMDIFVQIRKADTSSRNGT-----CODSSWLD 125

QY 64 --MIKER-----YMYLTEEILKENPNVCEYMAPSLDARQDMVVVEVPRLGKEAAVKAKE 116

Db 126 LRKIQERSGLGDETHGPEGLLVPPR--KTFAAARETEQVILGALENLFKNTNV----- 178

QY 117 WGPKSKITHLIIVCTTSGVDMPCADYQLTKLLGLRPVVKRYMMYQQXFGAGTVLR LAKD 176

Db 179 --NPKD-IGILVNSSWFNFTPSLSAMVNTFKLRNVRNFRNLGGMGCSAGVIAIDLAKD 235

QY 177 LAENNKGARVLVVCSEVTAATFRGSDTHLDSLVGOALFGDGAALIVGSDPVPETKEPI 236

Db 236 LLVHKNTVALVSTE--NITYNIYAGDNRMVSNCLFRVGGAAILLNKP-GDRRRSK 292

QY 237 FEMVWTAQTIAPDSEGAIDGHLR-----EAG-LTFHLLKDVPGI-----VSKNIT- 280

Db 293 YELVHTVRT-----HTGADDKSFRVCQQGDDENGKIGVSLSKDITDVAGRTVKKNIA TLGP 348

QY 281 -----KALVEAFEPGLGSDYNSIF--WIAHPGCPAILDQVEQKLAUK 320  
Db 349 LILPLSEKLLFFVTMGKLLFKDKIKHYVVPDFKLAIDHFCIHAGGRAVIDVLEKNLALA 408  
QY 321 PERONATREVLSEYGNMSSACVLFILDEMRKKS-TONGLKTTTGEGLWGLFGFG 374  
Db 409 PIDVEASRSTLHRFGNTSSSIWYELAYIEAKGRMKGNKV-----WQIALGSG 457

## RESULT 12

US-09-877-476-40

; Sequence 40, Application US/09877476

; Patent No. 6713664

; GENERAL INFORMATION:

; APPLICANT: Jaworski, Jan G.

; APPLICANT: Blacklock, Brenda J.

; TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA

; TITLE OF INVENTION: SYNTHASE POLYPEPTIDES

; FILE REFERENCE: 07148-108001

; CURRENT APPLICATION NUMBER: US/09/877,476

; CURRENT FILING DATE: 2001-06-08

; PRIOR APPLICATION NUMBER: US 60/210,326

; PRIOR FILING DATE: 2000-06-08

; NUMBER OF SEQ ID NOS: 56

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 40

; LENGTH: 506

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: 5' 173 amino acids from A. thaliana FAEL (SEQ ID

; OTHER INFORMATION: No:2) and 3' 333 amino acids from B. napus

; OTHER INFORMATION: elongase KCS (SEQ ID NO:4) having a mutation at

; OTHER INFORMATION: residue 307; designated At173 G307D; hypothetical

US-09-877-476-40

Query Match 9.0%; Score 180; DB 4; Length 506;

Best Local Similarity 23.2%; Pred. No. 3.3e-10; Indels 80; Gaps 17;

Matches 94; Conservative 62; Mismatches 170; Indels 80; Gaps 17;

QY 26 NPANCVEQSTY-PDFYFKITNSE-----HKTLEKEKFORMCDKS-----MIKRR-- 68

Db 75 NPVYLVDYSCYLPPLHRLVSVSKVMDIFVQIRKADTSSRNVACDDPSSLDPLRKIQERSG 134

QY 69 ---YMYLTEEILKENPNVCEYMAPSLDARQDMVVVEVPRLGKEAAVKAKEWQPKSKIT 125

Db 135 LGDETYSPEGLIHVPPR--KTFAASRETEKVIIGALENLFKNTNV-----NPKD-IG 184

QY 126 HLIIVCTTSGVDMPCADYQLTKLLGLRPVVKRYMMYQQXFGAGTVLR LAKLAENNKGAR 185

Db 185 ILVNVSSWFNFTPSLSAMVNTFKLRNVRNFRNLGGMGCSAGVIAIDLAKOLLHVHKNTY 244

QY 186 VLVVCEVTAATFRGSDTHLDSLVGOALFGDGAALIVGSDPVPETKEPIFEMVWTAQT 245

Db 245 ALVWSTE--NITYNIYAGDNRMVSNCLFRVGGAAILLNKP-GDRRRSKVELVHTVRT 301

QY 246 IAPDSEGAIDGHLR-----EAG-LTFHLLKDVPGI-----VSKNIT----- 280

Db 302 ---HTGADDKSFRVCQQGDDENGKIGVSLSKDITDVAGRTVKKNIA TLGPLILPLSEKL 357

QY 281 -----KALVEAFEPGLGSDYNSIF--WIAHPGCPAILDQVEQKLAKEPKMNAFRE 329

Db 358 LFFVTMGKLLFKDKIKHYVVPDFKLAIDHFCIHAGGRAVIDVLEKNLALAPIDVEASRS 417

QY 330 VLSEYGNMSSACVLFILDEMRKKS-TONGLKTTTGEGLWGLFGFG 374

Db 418 TLHRFGNTSSSIWYELAYIEAKGRMKGNKV-----WQIALGSG 457

## RESULT 13

US-08-657-749D-6

; Sequence 6, Application US/08657749D

Patent No. 6828475  
GENERAL INFORMATION:  
APPLICANT: METZ, JAMES G.  
APPLICANT: LARDIZABAL, KATHRYN D.  
APPLICANT: LASSNER, MICHAEL  
TITLE OF INVENTION: NUCLEIC ACID SEQUENCES ENCODING A PLANT CYTOPLASMIC  
TITLE OF INVENTION: PROTEIN INVOLVED IN FATTY ACYL-COA METABOLISM  
FILE REFERENCE: 16518.025  
CURRENT APPLICATION NUMBER: US/08/657,749D  
CURRENT FILING DATE: 1996-05-30  
PRIOR APPLICATION NUMBER: PCT/US94/13686  
PRIOR FILING DATE: 1994-11-30  
PRIOR APPLICATION NUMBER: 08/265,047  
PRIOR FILING DATE: 1994-06-23  
PRIOR APPLICATION NUMBER: 08/160,602  
PRIOR FILING DATE: 1993-11-30  
PRIOR APPLICATION NUMBER: 08/066,299  
PRIOR FILING DATE: 1993-05-20  
PRIOR APPLICATION NUMBER: PCT/US92/09863  
PRIOR FILING DATE: 1992-11-13  
PRIOR APPLICATION NUMBER: 07/933,411  
PRIOR FILING DATE: 1992-08-21  
PRIOR APPLICATION NUMBER: 07/796,256  
PRIOR FILING DATE: 1991-11-20  
NUMBER OF SEQ ID NOS: 53  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 6  
LENGTH: 521  
TYPE: PRT  
ORGANISM: Simmondsia chinensis  
US-08-657-749D-6

Query Match 9.0%; Score 179.5; DB 4; Length 521;  
Best Local Similarity 24.8%; Pred. No. 3.9e-10;  
Matches 93; Conservative 59; Mismatches 154; Indels 69; Gaps 15;  
QY 55 KFORMCDKSMIKRYMYLTBEILKENPNVCEYMAPSLDARQDMVVVEPRLGKEAAVKAI 114  
DB 146 EFORKILERAGMGRETVPESVTKVPE-----PSIAAAR-----AAEEMVYGAIDEVL 195  
QY 115 KENGQPKSKITHLIVCTTSVDMPGADYQTLKLLGRPYVKRYMMYQOGXFGAGTVLR 174  
DB 196 EKTGVKPKQIGILVNCNLFNPTPSLSSMIVNHYKLRNITLSYNLGCMGCSAGLISIDLA 255  
QY 175 KDLAENKNGARLVVCGSEVTAFTFRGSDTHLDSLVGOALFGDGAALIVGSDPVPEIEK 234  
DB 256 KDLQVTRNTYVLVSTENITNLYWYGNDRSM--LITNCLFRMGGAAILH--SRWRDRR 312  
QY 235 PIFEMVMTAQTAPDSEGAIDGHLR-----EAGLTFHLLKDVPGIVSK-----NIT 280  
DB 313 SKYQLLHTVRT-----HKGADKSYRCVQLQEDENNKVGVA--LSKDLMAVAGEALKANIT 366  
QY 281 -----KALV--BAFPLGSDYNSIFWIA-----HPGPAILLDOVEOK 316  
DB 367 TLGPLVLPMSQELLFATLVARVFKMTNVKYPIDPKLAARHFCIHAGGKAVLDELTN 426  
QY 317 LALKPEKMNATREIVSEYGNMSSACVLFILDEMRKKS--TQNGLKTGTGEGLEWGLFGFP 375  
DB 427 LEITPHLEPSRMTLVRFNGTSSSLWYELAYAEAKGRIRKGRD-----W--MIGFGS 478  
QY 376 GLTIETVW-----LRSV 387  
DB 479 GFKCNVWVWELASV 493

RESULT 14  
US-09-877-476-12  
Sequence 12, Application US/09877476  
Patent No. 6713664  
GENERAL INFORMATION:  
APPLICANT: Jaworski, Jan G.  
APPLICANT: Blacklock, Brenda J.  
TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA

TITLE OF INVENTION: SYNTHASE POLYPEPTIDES  
FILE REFERENCE: 07148-108001  
CURRENT APPLICATION NUMBER: US/09/877,476  
CURRENT FILING DATE: 2001-06-08  
PRIOR APPLICATION NUMBER: US 60/210,326  
PRIOR FILING DATE: 2000-06-08  
NUMBER OF SEQ ID NOS: 56  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 12  
LENGTH: 506  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: 5' 114 amino acids from A. thaliana FAE1 (SEQ ID  
NO:2) and 3' 392 amino acids from B. napus  
OTHER INFORMATION: elongase KCS (SEQ ID NO:4) having mutations at  
OTHER INFORMATION: residues 91 and 92; designated At114 L91C K92R  
US-09-877-476-12

Query Match 9.0%; Score 179; DB 4; Length 506;  
Best Local Similarity 22.2%; Pred. No. 4.2e-10;  
Matches 91; Conservative 63; Mismatches 155; Indels 100; Gaps 17;  
QY 27 PANC-VEQSTYDFYFKI-----TNSHKTELKKEFORMCDKS-----MIKR-----Y 69  
DB 88 PPHCRVSVSKVMDIFQIRKADTSRRNGT-----CDNSSWLDLRLKIQERSGLGDE 138  
QY 70 MYLTELKENPNVCEYMAPSLDARQDMVVVEPRLGKEAAVKAIKEWQPKSKITHLIV 129  
DB 139 THGPEGLQLVPPR--KTFAAAREETEQVIIGALENLFKNTNV-----NPKD-IGILVV 188  
QY 130 CTTSGVDMPGADYQTLKLLGRPYVKRYMMYQOGXFGAGTVLRDLAKLAEENKNGARLVV 189  
DB 189 NSSMFNPTSLSAMVNTFKLSNVRSFNLGCMGCSAGVIAIDLAKDLLHVHNTYALVV 248  
QY 190 CSEVTAFTFRGSDTHLDSLVGOALFGDGAALIVGSDPVPEIEKPIFEMVMTAQTAPD 249  
DB 249 STE--NITNYIAGDNRSMVSNCLFRVGGAAIILSNKP--GDRRSKYLVLHTVRT----- 301  
QY 250 SEGADGHLREAGLTFH-----LLKDVPGI-----VSKNIT----- 280  
DB 302 -----HTGADGKFRVCQGGDENGKIGVLSKDDITDVAGRTVKNTATLGLPLPLPS 354  
QY 281 -----KALVEAPEPLGSDYNSIF--WIAHPGGPAILLDOVEOKLAKPKEMNA 326  
DB 355 EXLLFPVTPMGKLLFKDKIKHYVVPDFKLAIDHFCIHAGGRAVIDVLEKNLALAPIDVEA 414  
QY 327 TREVLSEYGNMSSACVLFILDEMRKKS--TQNGLKTGTGEGLEWGLFGFG 374  
DB 415 SRSTLHRFGNTSSSIWYELAYIEAKRMKGNKV-----WQIALGSG 457

RESULT 15  
US-09-877-476-6  
Sequence 6, Application US/09877476  
Patent No. 6713664  
GENERAL INFORMATION:  
APPLICANT: Jaworski, Jan G.  
APPLICANT: Blacklock, Brenda J.  
TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA  
TITLE OF INVENTION: SYNTHASE POLYPEPTIDES  
FILE REFERENCE: 07148-108001  
CURRENT APPLICATION NUMBER: US/09/877,476  
CURRENT FILING DATE: 2001-06-08  
PRIOR APPLICATION NUMBER: US 60/210,326  
PRIOR FILING DATE: 2000-06-08  
NUMBER OF SEQ ID NOS: 56  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 6  
LENGTH: 505  
TYPE: PRT  
ORGANISM: Brassica napus  
US-09-877-476-6



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2005, 21:34:56 ; Search time 40 Seconds  
(without alignments)  
935.708 Million cell updates/sec

Title: US-10-031-918A-1  
Perfect score: 1997  
Sequence: 1 MVSVEIRKQRAEGPATIL.....LFGPGGLTIETVLRSAI 389

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1994	99.8	389	2 S35164	naringenin-chalcon
2	1969	98.6	389	2 S44370	naringenin-chalcon
3	1943	97.3	383	2 S35165	naringenin-chalcon
4	1938	97.2	389	2 S35166	naringenin-chalcon
5	1932	97.0	389	2 S26414	naringenin-chalcon
6	1932	96.7	389	2 S20932	naringenin-chalcon
7	1926	96.4	389	2 S35167	naringenin-chalcon
8	1926	96.4	389	2 S33210	naringenin-chalcon
9	1915	95.9	389	2 S20933	naringenin-chalcon
10	1911	95.7	389	2 S44367	naringenin-chalcon
11	1880	94.1	389	2 S35163	naringenin-chalcon
12	1875	93.9	370	2 S44368	naringenin-chalcon
13	1867	93.5	389	1 SYRUCP	naringenin-chalcon
14	1839	92.1	389	2 JQ2250	naringenin-chalcon
15	1818	91.0	389	2 S49202	naringenin-chalcon
16	1814	90.8	389	2 S49203	naringenin-chalcon
17	1811.5	90.7	388	1 SYSYC1	naringenin-chalcon
18	1809.5	90.6	388	2 JQ2249	naringenin-chalcon
19	1805.5	90.4	388	2 S60472	naringenin-chalcon
20	1800.5	90.2	388	1 SYSYC3	naringenin-chalcon
21	1800.5	90.2	388	2 S37098	naringenin-chalcon
22	1795.5	89.9	388	1 SYSYCN	naringenin-chalcon
23	1794.5	89.9	388	2 JQ2259	naringenin-chalcon
24	1767	88.5	389	1 SYRUCJ	naringenin-chalcon
25	1755	87.9	389	2 JCS136	naringenin-chalcon
26	1735	86.9	390	1 SYSKCD	naringenin-chalcon
27	1734	86.8	389	1 SYPUCN	naringenin-chalcon
28	1731	86.7	389	1 SYPUCA	naringenin-chalcon
29	1720	86.1	389	2 T07799	naringenin-chalcon

30	1705	85.4	396	2 S20515	naringenin-chalcon
31	1703.5	85.3	388	2 JC5516	naringenin-chalcon
32	1697	85.0	410	2 S12224	naringenin-chalcon
33	1695.5	84.9	394	1 SYJCCS	naringenin-chalcon
34	1680	84.1	398	2 S56699	naringenin-chalcon
35	1674.5	83.9	395	1 SYISCI	naringenin-chalcon
36	1667.5	83.5	395	1 SYISCI	naringenin-chalcon
37	1660.5	83.1	395	1 SYMUCN	naringenin-chalcon
38	1658	83.0	381	2 S12223	naringenin-chalcon
39	1652	82.7	398	2 S42523	naringenin-chalcon
40	1648	82.5	403	2 S55464	chalcone synthase
41	1637	82.0	391	2 T10713	naringenin-chalcon
42	1634	81.8	398	2 S58190	naringenin-chalcon
43	1632	81.7	400	1 SYZMCC	naringenin-chalcon
44	1614	80.8	398	2 S16275	naringenin-chalcon
45	1575	78.9	331	2 JQ1071	naringenin-chalcon

ALIGNMENTS

RESULT 1

S35164  
naringenin-chalcone synthase (EC 2.3.1.74) 2 - alfalfa  
N:Alternate names: chalcone synthase  
C:Species: Medicago sativa (alfalfa)  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
C:Accession: S35164  
R:Jungmans, H.; Dalkin, K.; Dixon, R.A.  
Plant Mol. Biol. 22, 239-253, 1993  
A:Title: Stress responses in alfalfa (Medicago sativa L.). 15. Characterization and exp.  
A:Reference number: S35163; MUID:93283629; PMID:8507827  
A:Accession: S35164  
A:Molecule type: mRNA  
A:Residues: 1-389 <JUN>  
A:Cross-references: UNIPROT:P30074; EMBL:L02902; NID:g166363; PIDN:AAA02824.1; PID:g166363  
C:Superfamily: Type III polyketide synthase  
C:Keywords: acyltransferase; coenzyme A; flavonoid biosynthesis

Query Match 99.8%; Score 1994; DB 2; Length 389;  
Best Local Similarity 99.7%; Pred. No. 4.7e-143;  
Matches 388; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	MVSVEIRKQRAEGPATILAI	GTANPANCVEQSTYDPFYFKIT	NSEHKTELKEKQRC	60
DB	1	MVSVEIRKQRAEGPATILAI	GTANPANCVEQSTYDPFYFKIT	NSEHKTELKEKQRC	60
QY	61	DKSMIKRRYMYLTETILKEN	PNVCYMAPSLDARDQMVVVE	VPRLGKEAAVKAKEWGQP	120
DB	61	DKSMIKRRYMYLTETILKEN	PNVCYMAPSLDARDQMVVVE	VPRLGKEAAVKAKEWGQP	120
QY	121	KSKITHLIVCTTSGVDMFG	ADYQLTKLGLRPYVKRYMY	QGGXHGATVPLAKDLAEN	180
DB	121	KSKITHLIVCTTSGVDMFG	ADYQLTKLGLRPYVKRYMY	QGGXHGATVPLAKDLAEN	180
QY	181	NKGARVLVVCSEVTATVTR	FGSDTHLDSLVGQALFGDGA	ALIVGSDPPEIEKPIFENV	240
DB	181	NKGARVLVVCSEVTATVTR	FGSDTHLDSLVGQALFGDGA	ALIVGSDPPEIEKPIFENV	240
QY	241	WTAQTIAPDSEGAIDGHLR	EAGLTATHLLKDVPGIVSK	NIITKALVEAFELGSDYNS	300
DB	241	WTAQTIAPDSEGAIDGHLR	EAGLTATHLLKDVPGIVSK	NIITKALVEAFELGSDYNS	300
QY	301	IHAHGGPAILDQVEQKAL	KEKNATREVLSEYGNMSAC	VCVLFILDEMRKSTQNGLKT	360
DB	301	IHAHGGPAILDQVEQKAL	KEKNATREVLSEYGNMSAC	VCVLFILDEMRKSTQNGLKT	360
QY	361	TGEGLEWGLFGFGPLGTI	ETVWLRSAI 389		
DB	361	TGEGLEWGLFGFGPLGTI	ETVWLRSAI 389		

RESULT 2

S44370  
naringenin-chalcone synthase (EC 2.3.1.74) - alfalfa  
C:Species: Medicago sativa (alfalfa)  
C>Date: 13-Jan-1995 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
C:Accession: S44370  
R:McKhann, H.I.; Hirsch, A.M.  
Plant Mol. Biol. 24, 767-777, 1994  
A:Title: Isolation of chalcone synthase and chalcone isomerase cDNAs from alfalfa (Medicago sativa L.)  
A:Reference number: S44370  
A:Accession: S44370  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-389 <MCK>  
A:Cross-references: UNIPROT:P30075; EMBL:U01021; NID:g393000; PIDN:AAB41559.1; PID:g393000  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1993  
C:Superfamily: Type III polyketide synthase  
C:Keywords: acyltransferase; coenzyme A

Query Match 98.6%; Score 1969; DB 2; Length 389;  
Best Local Similarity 98.5%; Pred. No. 3.6e-141;  
Matches 383; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 MVSVEIRKAQRAEGPATILAIQTANPANCVEQSTYDPDFYFKITNSEHKTTELKKEKFORMC 60  
DB 1 MVSVEIRKAQRAEGPATILAIQTANPANCVEQSTYDPDFYFKITNSEHKTTELKKEKFORMC 60  
QY 61 DKSMIKRRYMYLITEILKENPNVCEYMAPSLDARDQMVVVEVPRLGKEAAVKAKEWGQ 120  
DB 61 DKSMIKRRYMYLITEILKENPNVCEYMAPSLDARDQMVVVEVPRLGKEAAVKAKEWGQ 120  
QY 121 KSKITHLIVCTTSGVDMFGADYQTLTKLGLRPVYKRYMYQQGCFAGGTVLRLAKDLAEN 180  
DB 121 KSKITHLIVCTTSGVDMFGADYQTLTKLGLRPVYKRYMYQQGCFAGGTVLRLAKDLAEN 180  
QY 181 NKGARLVVVCSEVTAVTFRGPSDTHLDSLVGQALFGDGAALIVGSDPVPPEIEKPIFEMV 240  
DB 181 NKGARLVVVCSEVTAVTFRGPSDTHLDSLVGQALFGDGAALIVGSDPVPPEIEKPIFEMV 240  
QY 241 WTAQTAPSEGAIDGHLREAGLTFFHLLKDVPGIVSKNITKALVEAFEPGLGSDYNSIPW 300  
DB 241 WTAQTAPSEGAIDGHLREAGLTFFHLLKDVPGIVSKNITKALVEAFEPGLGSDYNSIPW 300  
QY 301 IAHPGGPAILDQVEQKALPKPKRNATREVLSEYGNMSSACVLFILDMRKKSTQNGLKT 360  
DB 301 IAHPGGPAILDQVEQKALPKPKRNATREVLSEYGNMSSACVLFILDMRKKSTQNGLKT 360  
QY 361 TGEGLBWGLVFGFGPGGLTIETVLRSAI 389  
DB 361 TGEGLBWGLVFGFGPGGLTIETVLRSAI 389

RESULT 3  
S35165  
naringenin-chalcone synthase (EC 2.3.1.74) 4 - alfalfa (fragment)  
N:Alternate names: chalcone synthase  
C:Species: Medicago sativa (alfalfa)  
C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 15-Mar-2004  
C:Accession: S35165; S36949  
R:Jungmans, H.; Dalkin, K.; Dixon, R.A.  
Plant Mol. Biol. 22, 239-253, 1993  
A:Title: Stress responses in alfalfa (Medicago sativa L.)  
A:Reference number: S35165  
A:Accession: S35165  
A:Molecule type: mRNA  
A:Residues: 1-331 <JUN>  
A:Cross-references: EMBL:L02903  
R:Jungmans, H.; Dalkin, K.; Dixon, R.A.  
submitted to the EMBL Data Library, September 1992  
A:Reference number: S36949  
A:Accession: S36949  
A:Molecule type: mRNA  
A:Residues: 53-383 <JU2>  
A:Cross-references: EMBL:L02903; NID:g166365; PIDN:AAA02825.1; PID:g166366

C:Superfamily: Type III polyketide synthase  
C:Keywords: acyltransferase; coenzyme A; flavonoid biosynthesis

Query Match 97.3%; Score 1943; DB 2; Length 383;  
Best Local Similarity 98.4%; Pred. No. 3.3e-139;  
Matches 377; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 7 IRKQARAGPATILAIQTANPANCVEQSTYDPDFYFKITNSEHKTTELKKEKFORMCDKSMIK 66  
DB 1 IRKQARAGPATILAIQTANPANCVEQSTYDPDFYFKITNSEHKTTELKKEKFORMCDKSMIK 60  
QY 67 RRYMYLITEILKENPNVCEYMAPSLDARDQMVVVEVPRLGKEAAVKAKEWGQPKSKITH 126  
DB 61 RRYMYLITEILKENPNVCEYMAPSLDARDQMVVVEVPRLGKEAAVKAKEWGQPKSKITH 120  
QY 127 LIVCTTSGVDMFGADYQTLTKLGLRPVYKRYMYQQGCFAGGTVLRLAKDLAENKKGARV 186  
DB 121 LIVCTTSGVDMFGADYQTLTKLGLRPVYKRYMYQQGCFAGGTVLRLAKDLAENKKGARV 180  
QY 187 LVVCSEVTAVTFRGPSDTHLDSLVGQALFGDGAALIVGSDPVPPEIEKPIFEMVMTAQT 246  
DB 181 LVVCSEVTAVTFRGPSDTHLDSLVGQALFGDGAALIVGSDPVPPEIEKPIFEMVMTAQT 240  
QY 247 APDSEGAIDGHLREAGLTFFHLLKDVPGIVSKNITKALVEAFEPGLGSDYNSIPWIAHPGG 306  
DB 241 APDSEGAIDGHLREAGLTFFHLLKDVPGIVSKNITKALVEAFEPGLGSDYNSIPWIAHPGG 300  
QY 307 PAILDQVEQKALPKPKRNATREVLSEYGNMSSACVLFILDMRKKSTQNGLKTGEGLE 366  
DB 301 PAILDQVEQKALPKPKRNATREVLSEYGNMSSACVLFILDMRKKSTQNGLKTGEGLE 360  
QY 367 WGLVFGFGPGGLTIETVLRSAI 389  
DB 361 WGLVFGFGPGGLTIETVLRSAI 383

RESULT 4  
S35166  
naringenin-chalcone synthase (EC 2.3.1.74) 8 - alfalfa  
N:Alternate names: chalcone synthase  
C:Species: Medicago sativa (alfalfa)  
C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
C:Accession: S35166  
R:Jungmans, H.; Dalkin, K.; Dixon, R.A.  
Plant Mol. Biol. 22, 239-253, 1993  
A:Title: Stress responses in alfalfa (Medicago sativa L.)  
A:Reference number: S35163; MUID:93283629; PMID:8507827  
A:Accession: S35166  
A:Molecule type: mRNA  
A:Residues: 1-389 <JUN>  
A:Cross-references: UNIPROT:P30076; EMBL:L02904; NID:g166367; PIDN:AAA02826.1; PID:g166368  
A:Note: the authors translated the codon GTT for residue 344 as Phe  
C:Superfamily: Type III polyketide synthase  
C:Keywords: acyltransferase; coenzyme A; flavonoid biosynthesis

Query Match 97.2%; Score 1941; DB 2; Length 389;  
Best Local Similarity 97.4%; Pred. No. 4.8e-139;  
Matches 379; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 MVSVEIRKAQRAEGPATILAIQTANPANCVEQSTYDPDFYFKITNSEHKTTELKKEKFORMC 60  
DB 1 MVSVEIRKAQRAEGPATILAIQTANPANCVEQSTYDPDFYFKITNSEHKTTELKKEKFORMC 60  
QY 61 DKSMIKRRYMYLITEILKENPNVCEYMAPSLDARDQMVVVEVPRLGKEAAVKAKEWGQ 120  
DB 61 DKSMIKRRYMYLITEILKENPNVCEYMAPSLDARDQMVVVEVPRLGKEAAVKAKEWGQ 120  
QY 121 KSKITHLIVCTTSGVDMFGADYQTLTKLGLRPVYKRYMYQQGCFAGGTVLRLAKDLAEN 180  
DB 121 KSKITHLIVCTTSGVDMFGADYQTLTKLGLRPVYKRYMYQQGCFAGGTVLRLAKDLAEN 180  
QY 181 NKGARLVVVCSEVTAVTFRGPSDTHLDSLVGQALFGDGAALIVGSDPVPPEIEKPIFEMV 240



Db 181 NKGARVLVVCSEVTAVTFRGSDTHLDSLVGQALFGDGAALIVGSDPVPPEIEKPIFEMV 240  
Qy 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFEPGLIGSDYNSIFW 300  
Db 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFEPGLIGSDYNSIFW 300  
Qy 301 IAHGPGPAILDQVEQKALPEKXNATREVLSEYGNMSSACVLFILDEMRRKSTQGLTKT 360  
Db 301 IAHGPGPAILDQVEQKALPEKXNATREVLSEYGNMSSACVLFILDEMRRKSTQGLTKT 360  
Qy 361 TGEGLWGVLFPGFGLTIETVLRVAI 389  
Db 361 TGEGLWGVLFPGFGLTIETVLRVAI 389

## RESULT 5

S26414  
naringenin-chalcone synthase (EC 2.3.1.74) - alfalfa  
C;Species: Medicago sativa (alfalfa)  
C;Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
C;Accession: S26414  
R;Buffard, D.; Breda, C.; El Turk, J.; Sallaud, C.; Kondorosi, A.; Esnault, R.  
submitted to the EMBL Data Library, August 1992  
A;Description: Molecular cloning of two chalcone synthase cDNA from alfalfa.  
A;Reference number: S26414  
A;Accession: S26414  
A;Molecule type: mRNA  
A;Residues: 1-389 <BUF>  
A;Cross-references: UNIPROT:P51078; EMBL:X69106; NID:g19590; PIDN:CAA48226.1; PID:g19591  
C;Superfamily: Type III polyketide synthase  
C;Keywords: acyltransferase; coenzyme A

Query Match 97.0%; Score 1938; DB 2; Length 389;  
Best Local Similarity 97.2%; Pred. No. 8e-139; 9; Indels 0; Gaps 0;  
Matches 378; Conservative 2; Mismatches 9;

Qy 1 MVSVEIRKAQRAEGPATILAI GTANPANCVEQSTYDPDFYFKITNSEHKTELKEKFORMC 60  
Db 1 MVSVEIRKAQRAEGPATILAI GTANPTNCVEQSTYDPDFYFKITNSEHKTELKEKFORMC 60  
Qy 61 DKSMIKRRYMYLTEILKENPNVCEYMAPSLDARDQMDVVEVPRLGKEAAVKAKEWGQP 120  
Db 61 DKSMIKRRYMYLTEILKENPSVCEIMAPSLDARDQMDVVEVPRLGKEAAVKAKEWGQP 120  
Qy 121 KSKITHLIVCTTSGVDMPGADYQLTLLGLRPYVKRYMYQQCXFAGGTVLRKDLAEN 180  
Db 121 KSKITHLIVCTTSGVDMPGADYQLTLLGLRPYVKRYMYQQCXFAGGTVLRKDLAEN 180  
Qy 181 NKGARVLVVCSEVTAVTFRGSDTHLDSLVGQALFGDGAALIVGSDPVPPEIEKPIFEMV 240  
Db 181 NKGARVLVVCSEVTAVTFRGSDTHLDSLVGQALFGDGAALIVGSDPVPPEIEKPIFEMV 240  
Qy 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFEPGLIGSDYNSIFW 300  
Db 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFEPGLIGSDYNSIFW 300  
Qy 301 IAHGPGPAILDQVEQKALPEKXNATREVLSEYGNMSSACVLFILDEMRRKSTQGLTKT 360  
Db 301 IAHGPGPAILDQVEQKALPEKXNATREVLSEYGNMSSACVLFILDEMRRKSTQGLTKT 360  
Qy 361 TGEGLWGVLFPGFGLTIETVLRVAI 389  
Db 361 TGEGLWGVLFPGFGLTIETVLRVAI 389

## RESULT 6

S20932  
naringenin-chalcone synthase (EC 2.3.1.74) 2 - garden pea  
N;Alternate names: chalcone synthase  
C;Species: Pisum sativum (garden pea)  
C;Date: 04-Dec-1992 #sequence\_revision 04-Dec-1992 #text\_change 09-Jul-2004  
C;Accession: S20932  
R;Ichinose, Y.; Kawamata, S.; Yamada, T.; An, C.; Kajiwara, T.; Shiraishi, T.; Oku, H.

Plant Mol. Biol. 18, 1009-1012, 1992  
A;Title: Molecular cloning of chalcone synthase cDNAs from Pisum sativum.  
A;Reference number: S20931; MUID:92256801; PMID:1581561  
A;Accession: S20932  
A;Molecule type: mRNA  
A;Residues: 1-389 <ICH>  
A;Cross-references: UNIPROT:Q01287; EMBL:X63334; NID:g20825; PIDN:CAA44934.1; PID:g20825  
C;Superfamily: Type III polyketide synthase  
C;Keywords: acyltransferase; coenzyme A; flavonoid biosynthesis

Query Match 96.7%; Score 1932; DB 2; Length 389;  
Best Local Similarity 96.4%; Pred. No. 2.3e-138;  
Matches 375; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

Qy 1 MVSVEIRKAQRAEGPATILAI GTANPANCVEQSTYDPDFYFKITNSEHKTELKEKFORMC 60  
Db 1 MVTVSEIRKAQRAEGPATILAI GTANPANCVEQSTYDPDFYFKITNSEHKTELKEKFORMC 60  
Qy 61 DKSMIKRRYMYLTEILKENPNVCEYMAPSLDARDQMDVVEVPRLGKEAAVKAKEWGQP 120  
Db 61 DKSMIKRRYMYLTEILKENPSLCEYMAPSLDARDQMDVVEVPRLGKEAAVKAKEWGQP 120  
Qy 121 KSKITHLIVCTTSGVDMPGADYQLTLLGLRPYVKRYMYQQCXFAGGTVLRKDLAEN 180  
Db 121 KSKITHLIVCTTSGVDMPGADYQLTLLGLRPYVKRYMYQQCXFAGGTVLRKDLAEN 180  
Qy 181 NKGARVLVVCSEVTAVTFRGSDTHLDSLVGQALFGDGAALIVGSDPVPPEIEKPIFEMV 240  
Db 181 NKGARVLVVCSEVTAVTFRGSDTHLDSLVGQALFGDGAALIVGSDPVPPEIEKPIFEMV 240  
Qy 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFEPGLIGSDYNSIFW 300  
Db 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFEPGLIGSDYNSIFW 300  
Qy 301 IAHGPGPAILDQVEQKALPEKXNATREVLSEYGNMSSACVLFILDEMRRKSTQGLTKT 360  
Db 301 IAHGPGPAILDQVEQKALPEKXNATREVLSEYGNMSSACVLFILDEMRRKSTQGLTKT 360  
Qy 361 TGEGLWGVLFPGFGLTIETVLRVAI 389  
Db 361 TGEGLWGVLFPGFGLTIETVLRVAI 389

## RESULT 7

S35167  
naringenin-chalcone synthase (EC 2.3.1.74) 9 - alfalfa  
N;Alternate names: chalcone synthase  
C;Species: Medicago sativa (alfalfa)  
C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
C;Accession: S35167  
R;Junghane, H.; Dalkin, K.; Dixon, R.A.  
Plant Mol. Biol. 22, 239-253, 1993  
A;Title: Stress responses in alfalfa (Medicago sativa L.). 15. Characterization and exp.  
A;Reference number: S35163; MUID:93283629; PMID:8507827  
A;Accession: S35167  
A;Molecule type: mRNA  
A;Residues: 1-389 <JUN>  
A;Cross-references: UNIPROT:P30077; EMBL:L02905; NID:g166369; PIDN:AAA02827.1; PID:g166369  
C;Superfamily: Type III polyketide synthase  
C;Keywords: acyltransferase; coenzyme A; flavonoid biosynthesis

Query Match 96.4%; Score 1926; DB 2; Length 389;  
Best Local Similarity 95.9%; Pred. No. 6.5e-138;  
Matches 373; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

Qy 1 MVSVEIRKAQRAEGPATILAI GTANPANCVEQSTYDPDFYFKITNSEHKTELKEKFORMC 60  
Db 1 MVSVEIRKAQRAEGPATILAI GTANPANCVEQSTYDPDFYFKITNSEHKTELKEKFORMC 60  
Qy 61 DKSMIKRRYMYLTEILKENPNVCEYMAPSLDARDQMDVVEVPRLGKEAAVKAKEWGQP 120  
Db 61 DKSMIKRRYMYLTEILKENPSVCEYMAPSLDARDQMDVVEVPRLGKEAAVKAKEWGQP 120

QY 121 KSKITHLIVCTTSGVDMPGADYQTLTKLGLRPVVKRYMMYQCGXFAAGTGVRLAKDLAEN 180  
DB 121 KSKITHLICTTSGVDMPGADYQTLTKLGLRPVVKRYMMYQCGXFAAGTGVRLAKDLAEN 180  
QY 181 NKGARVLVVCSEVAVTFRGSDTHLDSLVGQALFGDGAALIVGSDPVPPIEKPIFEMV 240  
DB 181 NKGARVLVVCSEVAVTFRGSDTHLDSLVGQALFGDGAALIVGSDPVPPIEKPIFEMV 240  
QY 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFPLGSDYNSIFW 300  
DB 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFPLGSDYNSIFW 300  
QY 301 IAHPGGPAILDQVEOKLALKEPKMATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
DB 301 IAHPGGPAILDQVEOKLALKEPKMATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
QY 361 TGEGLWGVLFPGFPGGLTIETVVLRSVAI 389  
DB 361 TGEGLWGVLFPGFPGGLTIETVVLHSVAI 389  
RESULT 8  
S33610  
naringenin-chalcone synthase (EC 2.3.1.74) 1 - garden pea  
N:Alternate names: chalcone synthase  
C:Species: Pisum sativum (garden pea)  
C>Date: 08-Dec-1993 #sequence\_revision 23-Feb-1996 #text\_change 09-Jul-2004  
C:Accession: S33610, S20931  
R:An, C.; Ichinose, Y.; Yamada, T.; Tanaka, Y.; Shiraishi, T.; Oku, H.  
Plant Mol. Biol. 21, 789-803, 1993  
A:Title: Organization of the genes encoding chalcone synthase in Pisum sativum.  
A:Reference number: S33610; MUID:93222478; PMID:8467077  
A:Accession: S33610  
A:Molecule type: DNA  
A:Residues: 1-389 <ANC>  
A:Cross-references: UNIPROT:Q01286; EMBL:D10661; NID:g391780; PIDN:BA041512.1; PID:g3917  
R:Ichinose, Y.; Kawamata, S.; Yamada, T.; An, C.; Kajiwara, T.; Shiraishi, T.; Oku, H.  
Plant Mol. Biol. 18, 1009-1012, 1992  
A:Title: Molecular cloning of chalcone synthase cDNAs from Pisum sativum.  
A:Reference number: S20931; MUID:92256801; PMID:1581561  
A:Accession: S20931  
A:Molecule type: mRNA  
A:Residues: 1-389 <ICH>  
A:Cross-references: EMBL:X63333; NID:g20823; PIDN:CAA44933.1; PID:g20824  
C:Superfamily: Type III polyketide synthase  
C:Genetics: 60/1  
C:Introns: 60/1  
C:Keywords: acyltransferase; coenzyme A; flavonoid biosynthesis  
Query Match 96.4%; Score 1926; DB 2; Length 389;  
Best Local Similarity 96.1%; Pred. No. 6 5e-138;  
Matches 374; Conservative 5; Mismatches 10; Indels 0; Gaps 0;  
QY 1 MVSVSIRKAQRAEGPATILAIQTANPANCVEQSTYDPFYFKITNSEHKTTELKEKFORMC 60  
DB 1 MVSVSIRKQRAEGPATILAIQTANPANCVEQSTYDPFYFKITNSEHKTTELKEKFORMC 60  
QY 61 DKSMIKRMYLITEILKENPNVCEYMAPSLDARDQDMVVEVPRLGKEAAVKAKEWGQP 120  
DB 61 DKSMIKRMYLITEILKENPNVCEYMAPSLDARDQDMVVEVPRLGKEAAVKAKEWGQP 120  
QY 121 KSKITHLIVCTTSGVDMPGADYQTLTKLGLRPVVKRYMMYQCGXFAAGTGVRLAKDLAEN 180  
DB 121 KSKITHLIVCTTSGVDMPGADYQTLTKLGLRPVVKRYMMYQCGXFAAGTGVRLAKDLAEN 180  
QY 181 NKGARVLVVCSEVAVTFRGSDTHLDSLVGQALFGDGAALIVGSDPVPPIEKPIFEMV 240  
DB 181 NKGARVLVVCSEVAVTFRGSDTHLDSLVGQALFGDGAALIVGSDPVPPIEKPIFEMV 240  
QY 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFPLGSDYNSIFW 300  
DB 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFPLGSDYNSIFW 300  
QY 301 IAHPGGPAILDQVEOKLALKEPKMATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
DB 301 IAHPGGPAILDQVEOKLALKEPKMATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
QY 361 TGEGLWGVLFPGFPGGLTIETVVLRSVAI 389  
DB 361 TGEGLWGVLFPGFPGGLTIETVVLHSVAI 389  
RESULT 10  
S44367  
naringenin-chalcone synthase (EC 2.3.1.74) - alfalfa  
C:Species: Medicago sativa (alfalfa)  
C>Date: 13-Jan-1995 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
C:Accession: S44367  
R:McKham, H.I.; Hirsch, A.M.  
Plant Mol. Biol. 24, 767-777, 1994  
A:Title: Isolation of chalcone synthase and chalcone isomerase cDNAs from alfalfa (Medic  
A:Reference number: S44367; MUID:94250839; PMID:8193301  
A:Accession: S44367  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: mRNA

QY 301 IAHPGGPAILDQVEOKLALKEPKMATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
DB 301 IAHPGGPAILDQVEOKLALKEPKMATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
QY 361 TGEGLWGVLFPGFPGGLTIETVVLRSVAI 389  
DB 361 TGEGLWGVLFPGFPGGLTIETVVLHSVAI 389  
RESULT 9  
S20933  
naringenin-chalcone synthase (EC 2.3.1.74) 3 - garden pea  
N:Alternate names: chalcone synthase  
C:Species: Pisum sativum (garden pea)  
C>Date: 04-Dec-1992 #sequence\_revision 04-Dec-1992 #text\_change 09-Jul-2004  
C:Accession: S20933  
R:Ichinose, Y.; Kawamata, S.; Yamada, T.; An, C.; Kajiwara, T.; Shiraishi, T.; Oku, H.  
Plant Mol. Biol. 18, 1009-1012, 1992  
A:Title: Molecular cloning of chalcone synthase cDNAs from Pisum sativum.  
A:Reference number: S20931; MUID:92256801; PMID:1581561  
A:Accession: S20933  
A:Molecule type: mRNA  
A:Residues: 1-389 <ICH>  
A:Cross-references: UNIPROT:Q01288; EMBL:X63335; NID:g20827; PIDN:CAA44935.1; PID:g20828  
C:Superfamily: Type III polyketide synthase  
C:Keywords: acyltransferase; coenzyme A; flavonoid biosynthesis  
Query Match 95.9%; Score 1915; DB 2; Length 389;  
Best Local Similarity 95.4%; Pred. No. 4 4e-137;  
Matches 371; Conservative 6; Mismatches 12; Indels 0; Gaps 0;  
QY 1 MVSVSIRKAQRAEGPATILAIQTANPANCVEQSTYDPFYFKITNSEHKTTELKEKFORMC 60  
DB 1 MVSVSIRKAQRAEGPATILAIQTANPANCVEQSTYDPFYFKITNSEHKTTELKEKFORMC 60  
QY 61 DKSMIKRMYLITEILKENPNVCEYMAPSLDARDQDMVVEVPRLGKEAAVKAKEWGQP 120  
DB 61 DKSMIKRMYLITEILKENPNVCEYMAPSLDARDQDMVVEVPRLGKEAAVKAKEWGQP 120  
QY 121 KSKITHLIVCTTSGVDMPGADYQTLTKLGLRPVVKRYMMYQCGXFAAGTGVRLAKDLAEN 180  
DB 121 KSKITHLIVCTTSGVDMPGADYQTLTKLGLRPVVKRYMMYQCGXFAAGTGVRLAKDLAEN 180  
QY 181 NKGARVLVVCSEVAVTFRGSDTHLDSLVGQALFGDGAALIVGSDPVPPIEKPIFEMV 240  
DB 181 NKGARVLVVCSEVAVTFRGSDTHLDSLVGQALFGDGAALIVGSDPVPPIEKPIFEMV 240  
QY 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFPLGSDYNSIFW 300  
DB 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFPLGSDYNSIFW 300  
QY 301 IAHPGGPAILDQVEOKLALKEPKMATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
DB 301 IAHPGGPAILDQVEOKLALKEPKMATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
QY 361 TGEGLWGVLFPGFPGGLTIETVVLRSVAI 389  
DB 361 TGEGLWGVLFPGFPGGLTIETVVLHSVAI 389  
RESULT 10  
S44367  
naringenin-chalcone synthase (EC 2.3.1.74) - alfalfa  
C:Species: Medicago sativa (alfalfa)  
C>Date: 13-Jan-1995 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
C:Accession: S44367  
R:McKham, H.I.; Hirsch, A.M.  
Plant Mol. Biol. 24, 767-777, 1994  
A:Title: Isolation of chalcone synthase and chalcone isomerase cDNAs from alfalfa (Medic  
A:Reference number: S44367; MUID:94250839; PMID:8193301  
A:Accession: S44367  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: mRNA

A;Residues: 1-389 <MCK>  
A;Cross-references: UNIPROT:P51077; EMBL:U01018; NID:G392994; PIDN:AAB41561.1; PID:G392994  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1993  
C;Superfamily: Type III polyketide synthase  
C;Keywords: acyltransferase; coenzyme A

Query Match 95.7%; Score 1911; DB 2; Length 389;  
Best Local Similarity 94.9%; Pred. No. 8.8e-137;  
Matches 363; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

QY 1 MVSVEIRKQRAEGPATILAIATGATPANCVEOSTYDPDFYFKITNSEHKTELKEKFORMC 60  
DB 1 MVSVEIRKQRAEGPATILAIATGATPANCVEOSTYDPDFYFKITNSEHKTELKEKFORMC 60

QY 61 DKSMIKRRYMYLVEELKENPNVCEYMAPSLDARQDMVVVEVPRLGKEAAVKAKEWGQP 120  
DB 61 DKSMIKRRYMYLVEELKENPNVCEYMAPSLDARQDMVVVEVPRLGKEAAVKAKEWGQP 120

QY 121 KSKITHLIVCTTSGVDMPGADYQTLKGLRPPYVKRYMYQQXFPAGGTVLRKOLAEN 180  
DB 121 KSKITHLIVCTTSGVDMPGADYQTLKGLRPPYVKRYMYQQXFPAGGTVLRKOLAEN 180

QY 181 NKGARVLVVCSEVTAVTFRGPSTHLDLSLVGQALFGDGAALIVGSDPVEIEKPIFEW 240  
DB 181 NKGARVLVVCSEVTAVTFRGPSTHLDLSLVGQALFGDGAALIVGSDPVEIEKPIFEW 240

QY 241 WTAQTIAPDSEGAIDGHLREAGLTFFHLLKDVPGIVSKNITKALVEAPEPLGISTDYNISIFW 300  
DB 241 WTAQTIAPDSEGAIDGHLREAGLTFFHLLKDVPGIVSKNITKALVEAPEPLGISTDYNISIFW 300

QY 301 IAHPPGPAILDQVEOKLAKPEKKNATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
DB 301 IAHPPGPAILDQVEOKLAKPEKKNATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360

QY 361 TGEGLWGVLFPGFGPLTIETVLRVAI 389  
DB 361 TGEGLWGVLFPGFGPLTIETVLRVAI 389

RESULT 11  
S35163  
naringenin-chalcone synthase (EC 2.3.1.74) 1 - alfalfa  
N;Alternate names: chalcone synthase  
C;Species: Medicago sativa (alfalfa)  
C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
C;Accession: S35163  
R;Jungmans, H.; Dalkin, K.; Dixon, R.A.  
Plant Mol. Biol. 22, 239-253, 1993  
A;Title: Stress responses in alfalfa (Medicago sativa L.). 15. Characterization and expression of the cDNA for the chalcone synthase gene  
A;Reference number: S35163; MUID:93283629; PMID:8507827  
A;Molecule type: mRNA  
A;Residues: 1-389 <JUN>  
A;Cross-references: UNIPROT:P30073; EMBL:L02901; NID:G166361; PIDN:AAA02823.1; PID:G166361  
A;Note: the authors translated the codon ATT for residue 284 as Val  
C;Superfamily: Type III polyketide synthase  
C;Keywords: acyltransferase; coenzyme A; flavonoid biosynthesis

Query Match 94.1%; Score 1880; DB 2; Length 389;  
Best Local Similarity 93.3%; Pred. No. 1.9e-134;  
Matches 363; Conservative 9; Mismatches 17; Indels 0; Gaps 0;

QY 1 MVSVEIRKQRAEGPATILAIATGATPANCVEOSTYDPDFYFKITNSEHKTELKEKFORMC 60  
DB 1 MVSVEIRKQRAEGPATILAIATGATPANCVEOSTYDPDFYFKITNSEHKTELKEKFORMC 60

QY 61 DKSMIKRRYMYLVEELKENPNVCEYMAPSLDARQDMVVVEVPRLGKEAAVKAKEWGQP 120  
DB 61 DKSMIKRRYMYLVEELKENPNVCEYMAPSLDARQDMVVVEVPRLGKEAAVKAKEWGQP 120

QY 121 KSKITHLIVCTTSGVDMPGADYQTLKGLRPPYVKRYMYQQXFPAGGTVLRKOLAEN 180  
DB 121 KSKITHLIVCTTSGVDMPGADYQTLKGLRPPYVKRYMYQQXFPAGGTVLRKOLAEN 180

QY 181 NKGARVLVVCSEVTAVTFRGPSTHLDLSLVGQALFGDGAALIVGSDPVEIEKPIFEW 240  
DB 181 NKGARVLVVCSEVTAVTFRGPSTHLDLSLVGQALFGDGAALIVGSDPVEIEKPIFEW 240

QY 241 WTAQTIAPDSEGAIDGHLREAGLTFFHLLKDVPGIVSKNITKALVEAPEPLGISTDYNISIFW 300  
DB 241 WTAQTIAPDSEGAIDGHLREAGLTFFHLLKDVPGIVSKNITKALVEAPEPLGISTDYNISIFW 300

QY 301 IAHPPGPAILDQVEOKLAKPEKKNATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
DB 301 IAHPPGPAILDQVEOKLAKPEKKNATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360

QY 361 TGEGLWGVLFPGFGPLTIETVLRVAI 389  
DB 361 TGEGLWGVLFPGFGPLTIETVLRVAI 389

RESULT 12  
S44368  
naringenin-chalcone synthase (EC 2.3.1.74) - alfalfa  
C;Species: Medicago sativa (alfalfa)  
C;Date: 13-Jan-1995 #sequence\_revision 10-Nov-1995 #text\_change 15-Mar-2004  
C;Accession: S44368  
R;McKhann, H.I.; Hirsch, A.M.  
Plant Mol. Biol. 24, 767-777, 1994  
A;Title: Isolation of chalcone synthase and chalcone isomerase cDNAs from alfalfa (Medicago sativa L.)  
A;Reference number: S44367; MUID:94250839; PMID:8193301  
A;Accession: S44368  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: mRNA  
A;Residues: 1-370 <MCK>  
A;Cross-references: EMBL:U01019; NID:G392996; PIDN:AAB41560.1; PID:G514830  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1993  
C;Superfamily: Type III polyketide synthase  
C;Keywords: acyltransferase; coenzyme A

Query Match 93.9%; Score 1875; DB 2; Length 370;  
Best Local Similarity 98.1%; Pred. No. 4.3e-134;  
Matches 363; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 20 LAIGTANPANCVEOSTYDPDFYFKITNSEHKTELKEKFORMCDKSMIKRRYMYLVEELKE 79  
DB 1 LAIGTANPANCVEOSTYDPDFYFKITNSEHKTELKEKFORMCDKSMIKRRYMYLVEELKE 79

QY 80 NPNVCEYMAPSLDARQDMVVVEVPRLGKEAAVKAKEWGQPKSKITHLIVCTTSGVDMPG 139  
DB 80 NPNVCEYMAPSLDARQDMVVVEVPRLGKEAAVKAKEWGQPKSKITHLIVCTTSGVDMPG 139

QY 140 ADYQTLKGLRPPYVKRYMYQQXFPAGGTVLRKOLAENNKGARVLVVCSEVTAVTFR 199  
DB 140 ADYQTLKGLRPPYVKRYMYQQXFPAGGTVLRKOLAENNKGARVLVVCSEVTAVTFR 199

QY 200 GPSDTHLDLSLVGQALFGDGAALIVGSDPVEIEKPIFEWMTAQTAPDSGAIDGHLR 259  
DB 200 GPSDTHLDLSLVGQALFGDGAALIVGSDPVEIEKPIFEWMTAQTAPDSGAIDGHLR 259

QY 260 EAGLTFHLLKDVPGIVSKNITKALVEAPEPLGISTDYNISIFWTAHPPGPAILDQVEOKLAL 319  
DB 260 EAGLTFHLLKDVPGIVSKNITKALVEAPEPLGISTDYNISIFWTAHPPGPAILDQVEOKLAL 319

QY 320 KPEKKNATREVLSEYGNMSSACVLFILDEMRRKSTONGLKTGEGLEWGVLFPGFGPLTI 379  
DB 320 KPEKKNATREVLSEYGNMSSACVLFILDEMRRKSTONGLKTGEGLEWGVLFPGFGPLTI 379

QY 380 ETVLRVAI 389  
DB 380 ETVLRVAI 389

RESULT 13  
SYFJCP  
naringenin-chalcone synthase (EC 2.3.1.74) I - kudzu vine

N;Alternate names: chalcone synthase I  
C;Species: *Pueraria lobata* (kudzu vine)  
C;Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004  
C;Accession: JQ0911; JC4539  
R;Nakaajima, O.; Akiyama, T.; Hakamatsuka, T.; Shibuya, M.; Noguchi, H.; Ebizuka, Y.; Sankawa, U.;  
submitted to JIPID, May 1991  
A;Reference number: JQ0911  
A;Accession: JQ0911  
A;Molecule type: DNA  
A;Residues: 1-389 <NAK>  
R;Cross-references: UNIPROT:P23569  
R;Nakaajima, O.; Shibuya, M.; Hakamatsuka, T.; Noguchi, H.; Ebizuka, Y.; Sankawa, U.  
Biol. Pharm. Bull. 19, 71-76, 1995  
A;Title: cDNA and genomic DNA clonings of chalcone synthase from *Pueraria lobata*.  
A;Reference number: JC4539; MUID:96418130; PMID:8820915  
A;Accession: JC4539  
A;Molecule type: mRNA  
A;Residues: 1-389 <NA2>  
C;Comment: This enzyme plays a central role in the biosynthesis of all classes of flavonoid compounds.  
C;Comment: This enzyme catalyzes the formation of chalcone from p-coumaroyl and malonyl  
C;regulatory point in isoflavonoidphytoalexin biosynthesis.  
C;Genetics:  
A;Gene: chs  
C;Superfamily: Type III polyketide synthase

Query Match	93.5%	Score 3867;	DB 1;	Length 389;
Best Local Similarity	91.0%;	Pred. No. 1.9e-133;		
Matches 354;	Conservative 20;	Mismatches 15;	Indels 0;	Gaps 0;
QY	1	MVSVSEIRKAQRAEGPATILGATNPANPCVQSTPYDFYFKITNSEHKTELEKEFORMC	60	
DB	1	MVSVAEIRKAQRAEGPATILGATNPANPCVQSTPYDFYFRTNSEHMTLEKEKFORMC	60	
QY	61	DKSMIKRRYMYLTBEILKENPNVCEYMAPSLDARQDMVVVEVPRLGKEAAVKAKEWGOP	120	
DB	61	DKSMIKRRYMYLTBEILKENPNMCAVMAPSLDARQDMVVVEVPLGKEAATAKKEWGOP	120	
QY	121	KSKITHLIVCTTSGVDMPGADYQLTKLGLRPVYKRYMYQQGXFAGGTIVLAKDLAEN	180	
DB	121	KSKITHLIFCTTSGVDMPGADYQLTQGLRPVYKRYMYQQCFAGGTIVLAKDLAEN	180	
QY	181	NGGARVLVVCSEVTAVTFRGPSDTHLDSLGQALFGDGAALIVGSDPVPEKEPIFENV	240	
DB	181	NGGARVLVVCSEITAVTFRGPSDTHLDSLSNGQALFGDGAAAVIVGSDPIQVPEKPLYELV	240	
QY	241	WTAQTIAPDPSGADIGHLRREAGLTTHLLKDVPGIVSKNITKALVEAFEPGLGTSYNSIFW	300	
DB	241	WTAQTIAPDPSGADIGHLRREVGLTTHLLKDVPGIVSKNIDKALFEAFNPLNITSYNSIFW	300	
QY	301	TAHPCGPAILQVQEQKALKPEKNATREVLSEYGNSSACVLFLIDEMRKSTQNGLKT	360	
DB	301	TAHPCGPAILQVQEQKGLPEKNMATRDVLSYGNSSACVLFLIDEMRKSANGLKT	360	
QY	361	TGEGLEWGVLFPGFGLTITFVILRSVAI	389	
DB	361	TGEGLEWGVLFPGFGLTITFVILRSVAI	389	

RESULT 14  
JQ2250  
naringenin-chalcone synthase (EC 2.3.1.74) - soybean  
N:Alternate names: chalcone synthase  
C:Species: Glycine max (soybean)  
C:Date: 19-May-1994 #sequence\_revision 19-May-1994 #text\_change 09-Jul-2004  
C:Accession: JQ2250  
R:Akada, S.; Kung, S.D.; Dube, S.K.  
Plant Physiol. 102, 321-323, 1993  
A:Title: Nucleotide sequence and putative regulatory elements of a nodule-development-s  
A:Reference number: JQ2250, MUID:94151429; PMID:8108501  
A:Accession: JQ2250  
A:Molecule type: DNA  
A:Residues: 1-389 >A>

A;Cross-references: UNIPROT:P30081; GB:M98871; NID:g169936; PIDN:AAA33950.1; PID:g169937  
C;Superfamily: Type III polyketide synthase  
C;Keywords: acyltransferase; coenzyme A; flavonoid biosynthesis

Query Match	92.1%;	Score 1839;	DB 2;	Length 389;
Best Local Similarity	90.8%;	Pred. No. 2.4e-131;		
Matches 351;	Conservative 19;	Mismatches 19;	Indels 0;	Gaps 0;

Qy	1	MVSVSIRKAQRAEGPATILAGTANPANCVSQSTVPDPFYFKITNSEHKTELKEKFORMC	60
		:     :     :     :     :     :     :     :     :     :	
Db	1	MVSAEIRQAQRAEGPATILAGTANPNRVDSQSTVPDPFYFRITNSDHWTELKEKFORMC	60
		:     :     :     :     :     :     :     :     :	
Qy	61	DKSMIKRRVMYITEILKENPNVCYMAPSLDARQDMVVVEVPRIGKGAAVKAIKEWGQP	120
		:     :     :     :     :     :     :     :     :	
Db	61	DKSMIKTRYMYLNEEILKENPNNCYMAPSLDARQDMVVVEVPLKGEAAVKAIKEWGQP	120
		:     :     :     :     :     :     :     :     :	
Qy	121	KSКИTHLIVCTTSGVDMPCADYQLTKLGLRPVVKRYMYQGXQCFAGGTVLRLAKDLAEN	180
		:     :     :     :     :     :     :     :     :	
Db	121	KSКИTHLIFCTTSGVDMFGADYQLTKLGLRPVVKRYMYQGCFCFAGGTVLRLAKDLAEN	180
		:     :     :     :     :     :     :     :     :	
Qy	181	NGGARVLVVCSEVTAVTFRGSDTHLDSLVGALFGDGAALIVGSDPVPETIEKPIFEVW	240
		:     :     :     :     :     :     :     :     :	
Db	181	NGGARVLVVCSEITAVTFRGSDTHLDSLVGALFGDGAALIVGSDPDPQVEKPLYLELV	240
		:     :     :     :     :     :     :     :     :	

```

Db      241  WTAQTAPDSEGAIDGHLREVGLTFHLLKDVPGITVSKNIDKALFEAFNPINISDYNISF 300
Qy      301  IAHGPGPAILDQVEQKLAKPEKNKATREVLSEYGNMSSACVLFILDEMRKKSSTONGLKT 360
Db      301  IAHGPGPAILDQVEQKLAKPEKNKATREVLSEYGNMSSACVLFILDEMRKKSSTONGLKT 360
Qy      361  TCGEGLEWGLVFGFGPGLTIETVVLRSVAI 389
Db      361  TCGEGLEWGLVFGFGPGLTIETVVLRSVAI 389

RESULT 15
S49202
naringenin-chalcone synthase (EC 2.3.1.74) - garden pea
C;Species: Pisum sativum (garden pea)
C;Date: 16-Feb-1995 #sequence_revision 12-May-1995 #text_change 09-Jul-2004
C;Accession: S49202
R;Hellens, R.P.
submitted to the EMBL Data Library, July 1994
A;Reference number: S49202
A;Accession: S49202
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-389 <HEL>
A;Cross-references: UNIPROT:P51081; EMBL:X80007; NID:g510542; PIDN:CAA56316.1;
C;Genetics:
A;Introns: 60/1
C;Superfamily: Type III polyketide synthase
C;Keywords: acyltransferase; coenzyme A

```

[illegible]

Db	181	NKGARVLVWCSEITAVTFRGPSTHLDLSLVGQALFGDGAADVIGSDPLDFVEKPLFELV	240
Qy	241	WTAQTIAPOSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFEPGLGSDYNSIFW	300
Db	241	WTAQTIPOSEGAIDGHLREAGLTFHLLKDVPSLVSKNIEKALVEAFQPLNISDYNSIFW	300
Qy	301	IAHPGGPAILDQVEQKALKPEKKNATREVLSEYGNMSSACVLFILDEMRRKKSSTONGLKT	360
Db	301	IAHPGGPAILDQVEAKLGLKQKQKQATRHVLSEYGNMSSACVLFILDEMRRKSKEDGLAT	360
Qy	361	TGEGLEWGVLFGEFGPGLTETVTLRSVA	388
Db	361	TGEGLEWGVLFGEFGPGLTETVTLHLSMA	388

Search completed: April 22, 2005, 21:45:05  
Job time : 44 secs

1103

608.01(A)

2163.07—  
(b)

2173.05(S)

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2005, 21:25:54 ; Search time 165 Seconds  
(without alignments)  
911.817 Million cell updates/sec

Title: US-10-031-918A-1

Perfect score: 1997  
Sequence: 1 MVSVEIRKAQRAEGPATIL.....LFGFGGLTFTVLRVAI 389

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: Geneseq1980s:\*
- 2: Geneseq1990s:\*
- 3: Geneseq2000s:\*
- 4: Geneseq2001s:\*
- 5: Geneseq2002s:\*
- 6: Geneseq2003as:\*
- 7: Geneseq2003bs:\*
- 8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1995	99.9	389	4	AAB68589 Chalcone
2	1994	99.8	389	5	ABG31832 Polyketid
3	1991	99.7	389	5	ABG31809 Polyketid
4	1991	99.7	389	5	ABG31804 Polyketid
5	1991	99.7	389	5	ABG31802 Polyketid
6	1991	99.7	389	5	ABG31817 Polyketid
7	1991	99.7	389	5	ABG31806 Polyketid
8	1990	99.6	389	5	ABG31821 Polyketid
9	1990	99.6	389	5	ABG31805 Polyketid
10	1990	99.6	389	5	ABG31813 Polyketid
11	1990	99.6	389	5	ABG31803 Polyketid
12	1990	99.6	389	5	ABG31812 Polyketid
13	1988	99.5	389	5	ABG31819 Polyketid
14	1988	99.5	389	5	ABG31808 Polyketid
15	1987	99.5	389	5	ABG31820 Polyketid
16	1987	99.5	389	5	ABG31822 Polyketid
17	1987	99.5	389	5	ABG31815 Polyketid
18	1986	99.4	389	5	ABG31807 Polyketid
19	1986	99.4	389	5	ABG31818 Polyketid
20	1986	99.4	389	5	ABG31801 Polyketid
21	1986	99.4	389	5	ABG31810 Polyketid
22	1986	99.4	389	5	ABG31811 Polyketid
23	1985	99.4	389	5	ABG31816 Polyketid
24	1969	98.6	389	8	ADM32822 Amino aci
25	1933	96.8	389	6	ABU94275 Trifolium

#### ALIGNMENTS

##### RESULT 1

AAB68589

ID AAB68589 standard; protein; 389 AA.

XX AC AAB68589;

XX DT 24-APR-2001 (first entry)

XX DE Chalcone synthase.

XX KW Chalcone synthase; enzyme; protein coordinate data; alfalfa;

XX KW polyketide synthase; plant phenylpropanoid biosynthesis; antibiotic;

XX KW anticancer agent; antifungal agent.

XX OS Medicago sativa.

XX FN WO200107579-A2.

XX PD 01-FEB-2001.

XX PF 27-JUL-2000; 2000WO-US020674.

XX PR 27-JUL-1999; 99US-0145898P.

XX (SALK ) SALK INST BIOLOGICAL STUDIES.

XX Noel JP, Ferrer J, Jez J, Austin M, Bowman M;

XX WPI; 2001-159712/16.

XX PT New isolated polyketide synthases, useful for identifying substrates and inhibitors and for producing potential pharmaceuticals.

XX PS Disclosure; Page 12; 213pp; English.

XX CC The present sequence is chalcone synthase (CHS) from Medicago sativa (alfalfa), which is a polyketide synthase. The chalcone synthase has at least 14 active site alpha-carbon atoms with structural coordinates defined in the specification. CHS plays an essential role in the biosynthesis of plant phenylpropanoids. Crystalline forms of CHS are used to determine the three-dimensional co-ordinates of polyketide synthases (PKS), especially of the active site, and from this information potential substrates and inhibitors of PKS can be modelled and tested, also predictions may be made about activity and substrate specificity of putative PKS. The co-ordinates may also be used to design mutant PKS that can produce different and/or new polyketide compounds, potentially useful as antibiotics, anticancer and antifungal agents

XX

SQ Sequence 389 AA;

Query Match 99.9%; Score 1995; DB 4; Length 389;  
 Best Local Similarity 100.0%; Pred. No. 6e-203;  
 Matches 389; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVSVEIRKAQRAEGPATILAI GTANPANCVEQSTYDPDFYFKITNSEHKTELKEKFORMC 60  
 DB 1 MVSVEIRKAQRAEGPATILAI GTANPANCVEQSTYDPDFYFKITNSEHKTELKEKFORMC 60

QY 61 DKSMIKRRYMYLITEILKENPNVCEYMAPSLDARDQMDVVEVPRLGKEAAVKAKEWGOP 120  
 DB 61 DKSMIKRRYMYLITEILKENPNVCEYMAPSLDARDQMDVVEVPRLGKEAAVKAKEWGOP 120

QY 121 KSKITHLIVCTTSGVDMPGADYQLT KLGRLPYVKRYMYQQGXFGAGTVLRLAKDLAEN 180  
 DB 121 KSKITHLIVCTTSGVDMPGADYQLT KLGRLPYVKRYMYQQGXFGAGTVLRLAKDLAEN 180

QY 181 NKGARVLVVCSEVTA VTFRGPSTHLDLSLVGOALFGDGAALIVGSDPVPETKEKPIFEMV 240  
 DB 181 NKGARVLVVCSEVTA VTFRGPSTHLDLSLVGOALFGDGAALIVGSDPVPETKEKPIFEMV 240

QY 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFELGISDYSNIFW 300  
 DB 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFELGISDYSNIFW 300

QY 301 IAHPGGPAILDQVEQKALKEPKMNATREVLSEYGNMSSACVLFILDEMRRKKSQTQGLKT 360  
 DB 301 IAHPGGPAILDQVEQKALKEPKMNATREVLSEYGNMSSACVLFILDEMRRKKSQTQGLKT 360

QY 361 TGEGLWGVLFPGFGLTIETVVLRSVAI 389  
 DB 361 TGEGLWGVLFPGFGLTIETVVLRSVAI 389

RESULT 2  
 ABG31832  
 ID ABG31832 standard; protein; 389 AA.  
 XX ABG31832;  
 XX 05-NOV-2002 (first entry)  
 DT Polyketide synthase.  
 DE Polyketide synthase.  
 XX Polyketide synthase; enzyme; protein coordinate data; PKS; resveratrol;  
 KW chalcone; crystal structure.  
 XX Mammalia.  
 XX WO200257418-A2.  
 XX 25-JUL-2002.  
 XX 14-DEC-2001; 2001WO-US048523.  
 XX 15-DEC-2000; 2000US-0255811P.  
 XX (SALK ) SALK INST BIOLOGICAL STUDIES.  
 XX Noel. JP, Austin MB, Bowman ME;  
 XX WPI; 2002-590730/63.  
 XX Producing mutant polyketide synthase (PKS) by comparing crystal structure  
 PT of wild-type (wt) PKS with crystal structure of second PKS and  
 PT substituting amino acids of wt PKS with amino acids at homologous  
 PT positions in second PKS.  
 XX Claim 2; Page 13; 243pp; English.  
 PS The invention relates to a method of producing mutant polyketide synthase  
 CC (PKS) comprising: (a) comparing a crystal structure of a wild-type PKS

CC with a crystal structure of a second PKS; (b) substituting one or more  
 CC amino acids of the wild-type PKS with the amino acid residues at  
 CC homologous positions in the second PKS; and (c) producing the mutant PKS.  
 CC The method is useful for producing a mutant polyketide synthase, and for  
 CC altering the activity of PKS, where the altered activity results in the  
 CC formation of the product of the second PKS instead of the product of the  
 CC wild-type PKS. Preferably, the altered activity results in the formation  
 CC of resveratrol instead of chalcone or in the formation of both  
 CC resveratrol instead of chalcone. The crystalline form of PKS is useful  
 CC for determining the position of specific alpha-carbon atoms and R-groups  
 CC associated with it, comprising the active site, in three-dimensional  
 CC space. It is also suitable for X-ray or neutron diffraction analysis to  
 CC determine three-dimensional structure of mutant PKS and to design  
 CC additional mutants. The crystallisation can serve as further  
 CC purification. Because the synthase may crystallise in more than one  
 CC crystal form, the structural coordinates of alpha-carbon atoms of an  
 CC active site determined from a synthase or its portions are useful to  
 CC solve structure of other crystal forms of synthases. The present sequence  
 CC represents wild-type PKS amino acid sequence used to make the PKS mutants  
 CC of the invention  
 XX

SQ Sequence 389 AA;

Query Match 99.8%; Score 1994; DB 5; Length 389;  
 Best Local Similarity 99.7%; Pred. No. 7.7e-203;  
 Matches 388; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MVSVEIRKAQRAEGPATILAI GTANPANCVEQSTYDPDFYFKITNSEHKTELKEKFORMC 60  
 DB 1 MVSVEIRKAQRAEGPATILAI GTANPANCVEQSTYDPDFYFKITNSEHKTELKEKFORMC 60

QY 61 DKSMIKRRYMYLITEILKENPNVCEYMAPSLDARDQMDVVEVPRLGKEAAVKAKEWGOP 120  
 DB 61 DKSMIKRRYMYLITEILKENPNVCEYMAPSLDARDQMDVVEVPRLGKEAAVKAKEWGOP 120

QY 121 KSKITHLIVCTTSGVDMPGADYQLT KLGRLPYVKRYMYQQGXFGAGTVLRLAKDLAEN 180  
 DB 121 KSKITHLIVCTTSGVDMPGADYQLT KLGRLPYVKRYMYQQGXFGAGTVLRLAKDLAEN 180

QY 181 NKGARVLVVCSEVTA VTFRGPSTHLDLSLVGOALFGDGAALIVGSDPVPETKEKPIFEMV 240  
 DB 181 NKGARVLVVCSEVTA VTFRGPSTHLDLSLVGOALFGDGAALIVGSDPVPETKEKPIFEMV 240

QY 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFELGISDYSNIFW 300  
 DB 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFELGISDYSNIFW 300

QY 301 IAHPGGPAILDQVEQKALKEPKMNATREVLSEYGNMSSACVLFILDEMRRKKSQTQGLKT 360  
 DB 301 IAHPGGPAILDQVEQKALKEPKMNATREVLSEYGNMSSACVLFILDEMRRKKSQTQGLKT 360

QY 361 TGEGLWGVLFPGFGLTIETVVLRSVAI 389  
 DB 361 TGEGLWGVLFPGFGLTIETVVLRSVAI 389

RESULT 3  
 ABG31809  
 ID ABG31809 standard; protein; 389 AA.  
 XX ABG31809;  
 XX 05-NOV-2002 (first entry)  
 DT Polyketide synthase mutant #9.  
 XX Polyketide synthase; enzyme; protein coordinate data; PKS; resveratrol;  
 KW chalcone; crystal structure; mutant; muten.  
 XX Mammalia.  
 OS Synthetic.  
 XX Key Location/Qualifiers  
 FH







KW chalcone; crystal structure; mutant; mutein.  
XX Mammalia.  
OS Synthetic.  
XX Key Location/Qualifiers  
FH Misc-difference 258 /note= "Wild type Leu substituted by Val"  
FT WT0200257418-A2.  
XX PN  
XX PD  
XX 25-JUL-2002.  
XX 14-DEC-2001; 2001WO-US048523.  
XX 15-DEC-2000; 2000US-0255811P.  
XX (SALK ) SALK INST BIOLOGICAL STUDIES.  
XX PI Noel JP, Austin MB, Bowman ME;  
XX WPI; 2002-590730/63.  
XX Producing mutant polyketide synthase (PKS) by comparing crystal structure of wild-type (wt) PKS with crystal structure of second PKS and substituting amino acids of wt PKS with amino acids at homologous positions in second PKS.  
XX Claim 4; Page; 243pp; English.  
XX The invention relates to a method of producing mutant polyketide synthase (PKS) comprising: (a) comparing a crystal structure of a wild-type PKS with a crystal structure of a second PKS; (b) substituting one or more amino acids of the wild-type PKS with the amino acid residues at homologous positions in the second PKS; and (c) producing the mutant PKS. The method is useful for producing a mutant polyketide synthase, and for altering the activity of PKS, where the altered activity results in the formation of the product of the second PKS instead of the product of the wild-type PKS. Preferably, the altered activity results in the formation of resveratrol instead of chalcone. The crystalline form of PKS is useful for determining the position of specific alpha-carbon atoms and R-groups associated with it, comprising the active site, in three-dimensional space. It is also suitable for X-ray or neutron diffraction analysis to determine three-dimensional structure of mutant PKS and to design additional mutants. The crystallisation can serve as further purification. Because the synthase may crystallise in more than one crystal form, the structural coordinates of alpha-carbon atoms of an

Query Match  
Best Local Similarity 99.7%; Score 1991; DB 5; Length 389;  
Matches 387; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MVSVSIRKQAREGPATILAIATNPANCVESQSTYPPDFFKITNSEHKTLEKEQRM 60  
DB 1 MVSVSIRKQAREGPATILAIATNPANCVESQSTYPPDFFKITNSEHKTLEKEQRM 60  
QY 61 DKSMIKRMYLVEITLKENPNVCEYMAPSLDARDQMWVVEVPRLGEAAVKAKEWG 120  
DB 61 DKSMIKRMYLVEITLKENPNVCEYMAPSLDARDQMWVVEVPRLGEAAVKAKEWG 120  
QY 121 KSKITHLVCTTSGVDMPCGADYQLTKLGLRPVYKRYMYQQGXFAGTGLRLAKDLAEN 180  
DB 121 KSKITHLVCTTSGVDMPCGADYQLTKLGLRPVYKRYMYQQGXFAGTGLRLAKDLAEN 180  
QY 181 NKGARVLVVCSEVTAVTFRGPSTHLDLSLVGQALFGDGAALIVGSDPVEIEKPIFEMV 240

DB 181 NKGARVLVVCSEVTAVTFRGPSTHLDLSLVGQALFGDGAALIVGSDPVEIEKPIFEMV 240  
QY 241 WTAQTIAPDSEGAIDGHLREAGLTFHLKLDVPGIVSKNITKALVEAFEPGLGIDYNSIFW 300  
DB 241 WTAQTIAPDSEGAIDGHLREAGLTFHLKLDVPGIVSKNITKALVEAFEPGLGIDYNSIFW 300  
QY 301 IAHPGGPAILDQVEQKALKPEKMNATREVLSEYGNMSACVLFILDEMRKKSQTONGLKT 360  
DB 301 IAHPGGPAILDQVEQKALKPEKMNATREVLSEYGNMSACVLFILDEMRKKSQTONGLKT 360  
QY 361 TGEGLWGVLFPGPGLTITETVLRVAI 389  
DB 361 TGEGLWGVLFPGPGLTITETVLRVAI 389  
RESULT 7  
ID ABG31806 standard; protein; 389 AA.  
XX AC ABG31806;  
XX 05-NOV-2002 (first entry)  
XX DE Polyketide synthase mutant #6.  
XX KW Polyketide synthase; enzyme; protein coordinate data; PKS; resveratrol;  
XX KM chalcone; crystal structure; mutant; mutein.  
XX OS Mammalia.  
XX OS Synthetic.  
XX FH Key Location/Qualifiers  
FT Misc-difference 133 /note= "Wild type Ser substituted by Thr"  
XX PN WO200257418-A2.  
XX PD 25-JUL-2002.  
XX PF 14-DEC-2001; 2001WO-US048523.  
XX PR 15-DEC-2000; 2000US-0255811P.  
XX PA (SALK ) SALK INST BIOLOGICAL STUDIES.  
XX PI Noel JP, Austin MB, Bowman ME;  
XX WPI; 2002-590730/63.  
XX Producing mutant polyketide synthase (PKS) by comparing crystal structure of wild-type (wt) PKS with crystal structure of second PKS and substituting amino acids of wt PKS with amino acids at homologous positions in second PKS.  
XX Claim 4; Page; 243pp; English.  
XX The invention relates to a method of producing mutant polyketide synthase (PKS) comprising: (a) comparing a crystal structure of a wild-type PKS with a crystal structure of a second PKS; (b) substituting one or more amino acids of the wild-type PKS with the amino acid residues at homologous positions in the second PKS; and (c) producing the mutant PKS. The method is useful for producing a mutant polyketide synthase, and for altering the activity of PKS, where the altered activity results in the formation of the product of the second PKS instead of the product of the wild-type PKS. Preferably, the altered activity results in the formation of resveratrol instead of chalcone. The crystalline form of PKS is useful for determining the position of specific alpha-carbon atoms and R-groups associated with it, comprising the active site, in three-dimensional space. It is also suitable for X-ray or neutron diffraction analysis to determine three-dimensional structure of mutant PKS and to design additional mutants. The crystallisation can serve as further purification. Because the synthase may crystallise in more than one crystal form, the structural coordinates of alpha-carbon atoms of an

CC active site determined from a synthase or its portions are useful to  
 CC solve structure of other crystal forms of synthases. The present sequence  
 CC represents a mutant polyketide synthase of the invention. Note: The  
 CC present sequence is not shown in the specification but was derived from  
 CC the wild-type PKS sequence given on page 13 (see ABG31832)  
 XX  
 SQ Sequence 389 AA;

Query Match 99.7%; Score 1991; DB 5; Length 389;  
 Best Local Similarity 99.5%; Pred. No. 1.6e-202;  
 Matches 387; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MVSVEIRKAQRAEGPATILAGTANPANCVEQSTYDPDFYFKITNSEHKTLEKEFORMC 60  
 DB 1 MVSVEIRKAQRAEGPATILAGTANPANCVEQSTYDPDFYFKITNSEHKTLEKEFORMC 60

QY 61 DKSMIKRMYLTYEELKENPNVCEYMAPSLDARQDMVVVEVPRLGKEAAVKAKEWGQP 120  
 DB 61 DKSMIKRMYLTYEELKENPNVCEYMAPSLDARQDMVVVEVPRLGKEAAVKAKEWGQP 120

QY 121 KSKITHLIVCTTSGVDMFGADYQTLKLLGLRPYVKRYMYQQGXFGAGTIVRLAKDLAEN 180  
 DB 121 KSKITHLIVCTTSGVDMFGADYQTLKLLGLRPYVKRYMYQQGXFGAGTIVRLAKDLAEN 180

QY 181 NKGARVLVVCSEVTAVTPRGSDTHLDSLVGOALFGDGAALIVGSDPVPPEIEKPIFEMV 240  
 DB 181 NKGARVLVVCSEVTAVTPRGSDTHLDSLVGOALFGDGAALIVGSDPVPPEIEKPIFEMV 240

QY 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVFGIVSKNITKALVEAFELGSDYNSIFW 300  
 DB 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVFGIVSKNITKALVEAFELGSDYNSIFW 300

QY 301 IAHPGGPAILDQVEQKALKEPKKNATREVLSEYGNMSSACVLFILDEMRKKSSTQNGLKT 360  
 DB 301 IAHPGGPAILDQVEQKALKEPKKNATREVLSEYGNMSSACVLFILDEMRKKSSTQNGLKT 360

QY 361 TGEGLWGVLFQFGPGLTIETVLRSAI 389  
 DB 361 TGEGLWGVLFQFGPGLTIETVLRSAI 389

RESULT 8  
 ABG31821  
 ID ABG31821 standard; protein; 389 AA.  
 AC ABG31821;  
 XX  
 DT 05-NOV-2002 (first entry)  
 XX  
 DE Polyketide synthase mutant #21.  
 XX  
 KW Polyketide synthase; enzyme; protein coordinate data; PKS; resveratrol;  
 KW chalcone; crystal structure; mutant; mutin.  
 XX  
 OS Mammalia.  
 OS Synthetic.  
 XX  
 FH \* Key Location/Qualifiers  
 FT Misc-difference 270  
 FT /note= "Wild type Asp substituted by Ala"  
 XX  
 PN WO200257418-A2.  
 XX  
 PD 25-JUL-2002.  
 XX  
 PF 14-DEC-2001; 2001WO-US048523.  
 XX  
 PR 15-DEC-2000; 2000US-0255811P.  
 XX  
 PA (SALK ) SALK INST BIOLOGICAL STUDIES.  
 XX  
 PI Noel JP, Austin MB, Bowman ME;  
 XX

DR WPI; 2002-590730/63.  
 XX Producing mutant polyketide synthase (PKS) by comparing crystal structure  
 FT of wild-type (wt) PKS with crystal structure of second PKS and  
 PT substituting amino acids of wt PKS with amino acids at homologous  
 PT positions in second PKS.  
 XX  
 XS Claim 4; Page; 243pp; English.

XX The invention relates to a method of producing mutant polyketide synthase  
 CC (PKS) comprising: (a) comparing a crystal structure of a wild-type PKS  
 CC with a crystal structure of a second PKS; (b) substituting one or more  
 CC amino acids of the wild-type PKS with the amino acid residues at  
 CC homologous positions in the second PKS; and (c) producing the mutant PKS.  
 CC The method is useful for producing a mutant polyketide synthase, and for  
 CC altering the activity of PKS, where the altered activity results in the  
 CC formation of the product of the second PKS instead of the product of the  
 CC wild-type PKS. Preferably, the altered activity results in the formation  
 CC of resveratrol instead of chalcone. The crystalline form of PKS is useful  
 CC for determining the position of specific alpha-carbon atoms and R-groups  
 CC associated with it, comprising the active site, in three-dimensional  
 CC space. It is also suitable for X-ray or neutron diffraction analysis to  
 CC determine three-dimensional structure of mutant PKS and to design  
 CC additional mutants. The crystallisation can serve as further  
 CC purification. Because the synthase may crystallise in more than one  
 CC crystal form, the structural coordinates of alpha-carbon atoms of an  
 CC active site determined from a synthase or its portions are useful to  
 CC solve structure of other crystal forms of synthases. The present sequence  
 CC represents a mutant polyketide synthase of the invention. Note: The  
 CC present sequence is not shown in the specification but was derived from  
 CC the wild-type PKS sequence given on page 13 (see ABG31832)  
 XX  
 SQ Sequence 389 AA;

Query Match 99.6%; Score 1990; DB 5; Length 389;  
 Best Local Similarity 99.5%; Pred. No. 2e-202;  
 Matches 387; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MVSVEIRKAQRAEGPATILAGTANPANCVEQSTYDPDFYFKITNSEHKTLEKEFORMC 60  
 DB 1 MVSVEIRKAQRAEGPATILAGTANPANCVEQSTYDPDFYFKITNSEHKTLEKEFORMC 60

QY 61 DKSMIKRMYLTYEELKENPNVCEYMAPSLDARQDMVVVEVPRLGKEAAVKAKEWGQP 120  
 DB 61 DKSMIKRMYLTYEELKENPNVCEYMAPSLDARQDMVVVEVPRLGKEAAVKAKEWGQP 120

QY 121 KSKITHLIVCTTSGVDMFGADYQTLKLLGLRPYVKRYMYQQGXFGAGTIVRLAKDLAEN 180  
 DB 121 KSKITHLIVCTTSGVDMFGADYQTLKLLGLRPYVKRYMYQQGXFGAGTIVRLAKDLAEN 180

QY 181 NKGARVLVVCSEVTAVTPRGSDTHLDSLVGOALFGDGAALIVGSDPVPPEIEKPIFEMV 240  
 DB 181 NKGARVLVVCSEVTAVTPRGSDTHLDSLVGOALFGDGAALIVGSDPVPPEIEKPIFEMV 240

QY 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVFGIVSKNITKALVEAFELGSDYNSIFW 300  
 DB 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVFGIVSKNITKALVEAFELGSDYNSIFW 300

QY 301 IAHPGGPAILDQVEQKALKEPKKNATREVLSEYGNMSSACVLFILDEMRKKSSTQNGLKT 360  
 DB 301 IAHPGGPAILDQVEQKALKEPKKNATREVLSEYGNMSSACVLFILDEMRKKSSTQNGLKT 360

QY 361 TGEGLWGVLFQFGPGLTIETVLRSAI 389  
 DB 361 TGEGLWGVLFQFGPGLTIETVLRSAI 389

RESULT 9  
 ABG31805  
 ID ABG31805 standard; protein; 389 AA.  
 XX  
 AC ABG31805;  
 XX

DT 05-NOV-2002 (first entry)  
XX Polyketide synthase mutant #5.  
XX Polyketide synthase; enzyme; protein coordinate data; PKS; resveratrol;  
KW chalcone; crystal structure; mutant; mutetin.  
XX Mammalia.  
OS Synthetic.  
XX Key Location/Qualifiers  
FH Misc-difference 131 /note= "Wild type Thr substituted by Ser"  
FT WT020257418-A2.  
XX XX  
XX 25-JUL-2002.  
XX 14-DEC-2001; 2001WO-US048523.  
XX 15-DEC-2000; 2000US-0255811P.  
XX (SALK ) SALK INST BIOLOGICAL STUDIES.  
XX Noel JP, Austin MB, Bowman ME;  
XX WPI; 2002-590730/63.  
XX Producing mutant polyketide synthase (PKS) by comparing crystal structure  
PT of wild-type (wt) PKS with crystal structure of second PKS and  
PT substituting amino acids of wt PKS with amino acids at homologous  
PT positions in second PKS.  
XX Claim 4; Page; 243pp; English.  
XX The invention relates to a method of producing mutant polyketide synthase  
CC (PKS) comprising: (a) comparing a crystal structure of a wild-type PKS  
CC with a crystal structure of a second PKS; (b) substituting one or more  
CC amino acids of the wild-type PKS with the amino acid residues at  
CC homologous positions in the second PKS; and (c) producing the mutant PKS.  
CC The method is useful for producing a mutant polyketide synthase, and for  
CC altering the activity of PKS, where the altered activity results in the  
CC formation of the product of the second PKS instead of the product of the  
CC wild-type PKS. Preferably, the altered activity results in the formation  
CC of resveratrol instead of chalcone. The crystalline form of PKS is useful  
CC for determining the position of specific alpha-carbon atoms and R-groups  
CC associated with it, comprising the active site, in three-dimensional  
CC space. It is also suitable for X-ray or neutron diffraction analysis to  
CC determine three-dimensional structure of mutant PKS and to design  
CC additional mutants. The crystallisation can serve as further  
CC purification. Because the synthase may crystallise in more than one  
CC crystal form, the structural coordinates of alpha-carbon atoms of an  
CC active site determined from a synthase or its portions are useful to  
CC solve structure of other crystal forms of synthases. The present sequence  
CC represents a mutant polyketide synthase of the invention. Note: The  
CC present sequence is not shown in the specification but was derived from  
CC the wild-type PKS sequence given on page 13 (see ABG31813)  
XX Sequence 389 AA;  
SQ

Query Match 99.6%; Score 1990; DB 5; Length 389;  
Best Local Similarity 99.5%; Pred. No. 2e-202;  
Matches 387; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MVSVSIRKQAEAGPATILAIATNPANCVEQSTYDPDFYFKITNSEHKTTELKEKQRM 60  
DB 1 MVSVSIRKQAEAGPATILAIATNPANCVEQSTYDPDFYFKITNSEHKTTELKEKQRM 60  
QY 61 DKSMIKRRYMYLTELKENPNVCYMAPSLDARQDMVVVEVPRLGEAAVKAKEWGQP 120  
DB 61 DKSMIKRRYMYLTELKENPNVCYMAPSLDARQDMVVVEVPRLGEAAVKAKEWGQP 120  
QY 121 KSKITHLIVCTTSGVDMPGADYQLTKLLGLRPYVKRYMMYQQGCFAGGTVLRKDLAEN 180

DB 121 KSKITHLIVCTTSGVDMPGADYQLTKLLGLRPYVKRYMMYQQGCFAGGTVLRKDLAEN 180  
QY 181 NKGARVLVVCSEVTAATFRGSDTHLDSLVGOALFGDGAALIVGSDPVPPIEKPIFEMV 240  
DB 181 NKGARVLVVCSEVTAATFRGSDTHLDSLVGOALFGDGAALIVGSDPVPPIEKPIFEMV 240  
QY 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFEPILGSDYNSIFW 300  
DB 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFEPILGSDYNSIFW 300  
QY 301 IAHPGGPAILDQVEQKALPKPKMNATREVLSEYGNMSSACVLFILDEMRKSTQNLKT 360  
DB 301 IAHPGGPAILDQVEQKALPKPKMNATREVLSEYGNMSSACVLFILDEMRKSTQNLKT 360  
QY 361 TGELEWGVLFPGPGLTITETVLRSAI 389  
DB 361 TGELEWGVLFPGPGLTITETVLRSAI 389  
RESULT 10  
ABG31813  
ID ABG31813 standard; protein; 389 AA.  
XX AC ABG31813;  
XX 05-NOV-2002 (first entry)  
XX Polyketide synthase mutant #13.  
XX Polyketide synthase; enzyme; protein coordinate data; PKS; resveratrol;  
KW chalcone; crystal structure; mutant; mutetin.  
XX Mammalia.  
OS Synthetic.  
XX Key Location/Qualifiers  
FH Misc-difference 160 /note= "Wild type Tyr substituted by Phe"  
FT WT0200257418-A2.  
XX 25-JUL-2002.  
XX 14-DEC-2001; 2001WO-US048523.  
XX 15-DEC-2000; 2000US-0255811P.  
XX (SALK ) SALK INST BIOLOGICAL STUDIES.  
XX Noel JP, Austin MB, Bowman ME;  
XX WPI; 2002-590730/63.  
XX Producing mutant polyketide synthase (PKS) by comparing crystal structure  
PT of wild-type (wt) PKS with crystal structure of second PKS and  
PT substituting amino acids of wt PKS with amino acids at homologous  
PT positions in second PKS.  
XX Claim 4; Page; 243pp; English.  
XX The invention relates to a method of producing mutant polyketide synthase  
CC (PKS) comprising: (a) comparing a crystal structure of a wild-type PKS  
CC with a crystal structure of a second PKS; (b) substituting one or more  
CC amino acids of the wild-type PKS with the amino acid residues at  
CC homologous positions in the second PKS; and (c) producing the mutant PKS.  
CC The method is useful for producing a mutant polyketide synthase, and for  
CC altering the activity of PKS, where the altered activity results in the  
CC formation of the product of the second PKS instead of the product of the  
CC wild-type PKS. Preferably, the altered activity results in the formation  
CC of resveratrol instead of chalcone. The crystalline form of PKS is useful  
CC for determining the position of specific alpha-carbon atoms and R-groups  
CC associated with it, comprising the active site, in three-dimensional  
CC space. It is also suitable for X-ray or neutron diffraction analysis to  
CC determine three-dimensional structure of mutant PKS and to design  
CC additional mutants. The crystallisation can serve as further  
CC purification. Because the synthase may crystallise in more than one  
CC crystal form, the structural coordinates of alpha-carbon atoms of an  
CC active site determined from a synthase or its portions are useful to  
CC solve structure of other crystal forms of synthases. The present sequence  
CC represents a mutant polyketide synthase of the invention. Note: The  
CC present sequence is not shown in the specification but was derived from  
CC the wild-type PKS sequence given on page 13 (see ABG31813)  
XX Sequence 389 AA;  
SQ

CC space. It is also suitable for X-ray or neutron diffraction analysis to  
 CC determine three-dimensional structure of mutant PKS and to design  
 CC additional mutants. The crystallisation can serve as further  
 CC purification. Because the synthase may crystallise in more than one  
 CC crystal form, the structural coordinates of alpha-carbon atoms of an  
 CC active site determined from a synthase or its portions are useful to  
 CC solve structure of other crystal forms of synthases. The present sequence  
 CC represents a mutant polyketide synthase of the invention. Note: The  
 CC present sequence is not shown in the specification but was derived from  
 CC the wild-type PKS sequence given on page 13 (see ABG31832)  
 XX  
 SQ Sequence 389 AA;

Query Match 99.6%; Score 1990; DB 5; Length 389;  
 Best Local Similarity 99.5%; Pred. No. 2e-202;  
 Matches 387; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MVSVEIRKAQRAEGPATILAGTANPANCVEQSTYDPFYFKITNSEHKTLEKQFQMC 60  
 DB 1 MVSVEIRKAQRAEGPATILAGTANPANCVEQSTYDPFYFKITNSEHKTLEKQFQMC 60

QY 61 DKSMIKRRYMYLTBEILKENPNVCEYMAPSLDARQDMVVVEVPRLGKEAAVKAKEWGQP 120  
 DB 61 DKSMIKRRYMYLTBEILKENPNVCEYMAPSLDARQDMVVVEVPRLGKEAAVKAKEWGQP 120

QY 121 KSKITHLIVCTTSGVDMPGADYQTLKGLRYPYVRYMYQOQXFGAGTIVLRKDLAEN 180  
 DB 121 KSKITHLIVCTTSGVDMPGADYQTLKGLRYPYVRYMYQOQXFGAGTIVLRKDLAEN 180

QY 181 NGGARVLVVCSEVTAVTFRGSPTDHLDSLVGOALFGDGAALIVGSDVPVEIEKPIFEMV 240  
 DB 181 NGGARVLVVCSEVTAVTFRGSPTDHLDSLVGOALFGDGAALIVGSDVPVEIEKPIFEMV 240

QY 241 WTAQTIAPDSEGAIDGHLREAGLTFFHLKDVFGIVSKNITKALVEAFPLGSDYNSIFW 300  
 DB 241 WTAQTIAPDSEGAIDGHLREAGLTFFHLKDVFGIVSKNITKALVEAFPLGSDYNSIFW 300

QY 301 IAHPGGPAILDQVEQKALPKPEKNATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
 DB 301 IAHPGGPAILDQVEQKALPKPEKNATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360

QY 361 TGEGLWGVLFQFGPGLTIETVLRSAI 389  
 DB 361 TGEGLWGVLFQFGPGLTIETVLRSAI 389

## RESULT 11

ABG31803  
 ID ABG31803 standard; protein; 389 AA.  
 XX  
 AC ABG31803;  
 XX  
 DT 05-NOV-2002 (first entry)  
 XX  
 DE Polyketide synthase mutant #3.  
 XX  
 KW Polyketide synthase; enzyme; protein coordinate data; PKS; resveratrol;  
 XX chalcone; crystal structure; mutant; mutein.  
 XX  
 OS Mammalia.  
 XX Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 99 /note= "Wild type Val substituted by Ala"  
 FT  
 XX WO200257418-A2.  
 XX  
 XX 25-JUL-2002.  
 XX  
 XX 14-DEC-2001; 2001WO-US048523.  
 XX  
 XX 15-DEC-2000; 2000US-025581P.  
 PR

XX (SALK ) SALK INST BIOLOGICAL STUDIES.  
 PA Noel JP, Austin MB, Bowman ME;  
 PI WPI; 2002-590730/63.  
 XX  
 DR Producing mutant polyketide synthase (PKS) by comparing crystal structure  
 XX of wild-type (wt) PKS with crystal structure of second PKS and  
 PT substituting amino acids of wt PKS with amino acids at homologous  
 PT positions in second PKS.  
 XX  
 PS Claim 4; Page; 243pp; English.  
 XX  
 CC The invention relates to a method of producing mutant polyketide synthase  
 CC (PKS) comprising: (a) comparing a crystal structure of a wild-type PKS  
 CC with a crystal structure of a second PKS; (b) substituting one or more  
 CC amino acids of the wild-type PKS with the amino acid residues at  
 CC homologous positions in the second PKS; and (c) producing the mutant PKS.  
 CC The method is useful for producing a mutant polyketide synthase, and for  
 CC altering the activity of PKS, where the altered activity results in the  
 CC formation of the product of the second PKS instead of the product of the  
 CC wild-type PKS. Preferably, the altered activity results in the formation  
 CC of resveratrol instead of chalcone. The crystalline form of PKS is useful  
 CC for determining the position of specific alpha-carbon atoms and R-groups  
 CC associated with it, comprising the active site, in three-dimensional  
 CC space. It is also suitable for X-ray or neutron diffraction analysis to  
 CC determine three-dimensional structure of mutant PKS and to design  
 CC additional mutants. The crystallisation can serve as further  
 CC purification. Because the synthase may crystallise in more than one  
 CC crystal form, the structural coordinates of alpha-carbon atoms of an  
 CC active site determined from a synthase or its portions are useful to  
 CC solve structure of other crystal forms of synthases. The present sequence  
 CC represents a mutant polyketide synthase of the invention. Note: The  
 CC present sequence is not shown in the specification but was derived from  
 CC the wild-type PKS sequence given on page 13 (see ABG31832)  
 XX  
 SQ Sequence 389 AA;

Query Match 99.6%; Score 1990; DB 5; Length 389;  
 Best Local Similarity 99.5%; Pred. No. 2e-202;  
 Matches 387; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MVSVEIRKAQRAEGPATILAGTANPANCVEQSTYDPFYFKITNSEHKTLEKQFQMC 60  
 DB 1 MVSVEIRKAQRAEGPATILAGTANPANCVEQSTYDPFYFKITNSEHKTLEKQFQMC 60

QY 61 DKSMIKRRYMYLTBEILKENPNVCEYMAPSLDARQDMVVVEVPRLGKEAAVKAKEWGQP 120  
 DB 61 DKSMIKRRYMYLTBEILKENPNVCEYMAPSLDARQDMVVVEVPRLGKEAAVKAKEWGQP 120

QY 121 KSKITHLIVCTTSGVDMPGADYQTLKGLRYPYVRYMYQOQXFGAGTIVLRKDLAEN 180  
 DB 121 KSKITHLIVCTTSGVDMPGADYQTLKGLRYPYVRYMYQOQXFGAGTIVLRKDLAEN 180

QY 181 NGGARVLVVCSEVTAVTFRGSPTDHLDSLVGOALFGDGAALIVGSDVPVEIEKPIFEMV 240  
 DB 181 NGGARVLVVCSEVTAVTFRGSPTDHLDSLVGOALFGDGAALIVGSDVPVEIEKPIFEMV 240

QY 241 WTAQTIAPDSEGAIDGHLREAGLTFFHLKDVFGIVSKNITKALVEAFPLGSDYNSIFW 300  
 DB 241 WTAQTIAPDSEGAIDGHLREAGLTFFHLKDVFGIVSKNITKALVEAFPLGSDYNSIFW 300

QY 301 IAHPGGPAILDQVEQKALPKPEKNATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
 DB 301 IAHPGGPAILDQVEQKALPKPEKNATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360

QY 361 TGEGLWGVLFQFGPGLTIETVLRSAI 389  
 DB 361 TGEGLWGVLFQFGPGLTIETVLRSAI 389

RESULT 12

```

ABG31812
ID ABG31812 standard; protein; 389 AA.
XX AC
XX ABG31812;
XX DT
XX 05-NOV-2002 (first entry)
XX DE
XX Polyketide synthase mutant #12.
XX KW
XX Polyketide synthase; enzyme; protein coordinate data; PKS; resveratrol;
XX chalcone; crystal structure; mutant; mutein.
XX KW
XX Mammalia.
XX OS
XX Synthetic.
XX FH
XX Key Location/Qualifiers
XX FT Misc-difference 159 /note= "Wild type Met substituted by Val"
XX PN
XX WO200257418-A2.
XX XX
XX 25-JUL-2002.
XX PF 14-DEC-2001; 2001WO-US048523.
XX PR 15-DEC-2000; 2000US-0255811P.
XX PA (SALK ) SALK INST BIOLOGICAL STUDIES.
XX PI Noel JP, Austin MB, Bowman ME;
XX WPI; 2002-590730/63.
XX DR
XX Producing mutant polyketide synthase (PKS) by comparing crystal structure
XX of wild-type (wt) PKS with crystal structure of second PKS and
XX substituting amino acids of wt PKS with amino acids at homologous
XX positions in second PKS.
XX PS
XX Claim 4; Page; 243pp; English.
XX CC
XX The invention relates to a method of producing mutant polyketide synthase
XX (PKS) comprising: (a) comparing a crystal structure of a wild-type PKS
XX with a crystal structure of a second PKS; (b) substituting one or more
XX amino acids of the wild-type PKS with the amino acid residues at
XX homologous positions in the second PKS; and (c) producing the mutant PKS.
XX The method is useful for producing a mutant polyketide synthase, and for
XX altering the activity of PKS, where the altered activity results in the
XX formation of the product of the second PKS instead of the product of the
XX wild-type PKS. Preferably, the altered activity results in the formation
XX of resveratrol instead of chalcone. The crystalline form of PKS is useful
XX for determining the position of specific alpha-carbon atoms and R-groups
XX associated with it, comprising the active site, in three-dimensional
XX space. It is also suitable for X-ray or neutron diffraction analysis to
XX determine three-dimensional structure of mutant PKS and to design
XX additional mutants. The crystallisation can serve as further
XX purification. Because the synthase may crystallise in more than one
XX crystal form, the structural coordinates of alpha-carbon atoms of an
XX active site determined from a synthase or its portions are useful to
XX solve structure of other crystal forms of synthases. The present sequence
XX represents a mutant polyketide synthase of the invention. Note: The
XX present sequence is not shown in the specification but was derived from
XX the wild-type PKS sequence given on page 13 (see ABG31832)
XX SQ
XX Sequence 389 AA;

Query Match 99.6%; Score 1990; DB 5; Length 389;
Best Local Similarity 99.5%; Pred. No. 2e-202;
Matches 387; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MVSVSIRKQRAEGPATILAIATNPANCVCEQSTYPDFYFKITNSEHKTELKEKQRM 60
DB 1 MVSVSIRKQRAEGPATILAIATNPANCVCEQSTYPDFYFKITNSEHKTELKEKQRM 60

```

CC formation of the product of the second PKS instead of the product of the  
 CC wild-type PKS. Preferably, the altered activity results in the formation  
 CC of resveratrol instead of chalcone. The crystalline form of PKS is useful  
 CC for determining the position of specific alpha-carbon atoms and R-groups  
 CC associated with it, comprising the active site, in three-dimensional  
 CC space. It is also suitable for X-ray or neutron diffraction analysis to  
 CC determine three-dimensional structure of mutant PKS and to design  
 CC additional mutants. The crystallisation can serve as further  
 CC purification. Because the synthase may crystallise in more than one  
 CC crystal form, the structural coordinates of alpha-carbon atoms of an  
 CC active site determined from a synthase or its portions are useful to  
 CC solve structure of other crystal forms of synthases. The present sequence  
 CC represents a mutant polyketide synthase of the invention. Note: The  
 CC present sequence is not shown in the specification but was derived from  
 CC the wild-type PKS sequence given on page 13 (see ABG31832)

XX SQ Sequence 389 AA;

Query Match 99.5%; Score 1988; DB 5; Length 389;  
 Best Local Similarity 99.5%; Pred. No. 3.3e-202;  
 Matches 387; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MVSVEIRKAQRAEGPATILAIQTANPANCVEQSTYDFYFKITNSEHKTLEKQFQRC 60  
 DB 1 MVSVEIRKAQRAEGPATILAIQTANPANCVEQSTYDFYFKITNSEHKTLEKQFQRC 60

QY 61 DKSMIKRMYLITEILKENPNVCEYMAPSLDARQDMVVEVPRIGKEAAVKAKEWGQP 120  
 DB 61 DKSMIKRMYLITEILKENPNVCEYMAPSLDARQDMVVEVPRIGKEAAVKAKEWGQP 120

QY 121 KSKITHLIVCTTSGVDMFGADYQTLKGLRPVVKRYMYQQGXFAGGTVLRLAKDLAEN 180  
 DB 121 KSKITHLIVCTTSGVDMFGADYQTLKGLRPVVKRYMYQQGXFAGGTVLRLAKDLAEN 180

QY 181 NKGARLVVCSVTAFTFRGSDTHLDSLVGQALFGDGAALIVGSDPVPPEIEKPIFENV 240  
 DB 181 NKGARLVVCSVTAFTFRGSDTHLDSLVGQALFGDGAALIVGSDPVPPEIEKPIFENV 240

QY 241 WTAQTIAPDSEGAIDGHLREAGLTFLHLKDVPGIVSKNITKALVEAFEPGLGSDYNSIFW 300  
 DB 241 WTAQTIAPDSEGAIDGHLREAGLTFLHLKDVPGIVSKNITKALVEAFEPGLGSDYNSIFW 300

QY 301 IAHPCGPAILDQVEOKLALKEPKMNATREVLSEYGNMSSACVLFILDEMRKKSSTONGLKT 360  
 DB 301 IAHPCGPAILDQVEOKLALKEPKMNATREVLSEYGNMSSACVLFILDEMRKKSSTONGLKT 360

QY 361 TGEGLWGVLFPGFGLTIETVLRSAI 389

# RESULT 14

ABG31808  
 ID ABG31808 standard; protein; 389 AA.  
 XX AC ABG31808;  
 XX 05-NOV-2002 (first entry)  
 XX Polyketide synthase mutant #8.  
 KW Polyketide synthase; enzyme; protein coordinate data; PKS; resveratrol;  
 KW chalcone; crystal structure; mutant; mutin.  
 XX Mammalia.  
 OS Synthetic.  
 XX Key Location/Qualifiers  
 FT Misc-difference 135 /note= "Wild type Val substituted by Pro"  
 XX WO200257418-A2.

PD 25-JUL-2002.  
 XX 14-DEC-2001; 2001WO-US048523.  
 XX 15-DEC-2000; 2000US-0255811P.  
 XX (SALK ) SALK INST BIOLOGICAL STUDIES.

XX Noel JP, Austin MB, Bowman MB;  
 WIPI; 2002-590730/63.

PT Producing mutant polyketide synthase (PKS) by comparing crystal structure  
 of wild-type (wt) PKS with crystal structure of second PKS and  
 PT substituting amino acids of wt PKS with amino acids at homologous  
 PT positions in second PKS.

XX Claim 4; Page; 243pp; English.

XX The invention relates to a method of producing mutant polyketide synthase  
 CC (PKS) comprising: (a) comparing a crystal structure of a wild-type PKS  
 CC with a crystal structure of a second PKS; (b) substituting one or more  
 CC amino acids of the wild-type PKS with the amino acid residues at  
 CC homologous positions in the second PKS; and (c) producing the mutant PKS.  
 CC The method is useful for producing a mutant polyketide synthase, and for  
 CC altering the activity of PKS, where the altered activity results in the  
 CC formation of the product of the second PKS instead of the product of the  
 CC wild-type PKS. Preferably, the altered activity results in the formation  
 CC of resveratrol instead of chalcone. The crystalline form of PKS is useful  
 CC for determining the position of specific alpha-carbon atoms and R-groups  
 CC associated with it, comprising the active site, in three-dimensional  
 CC space. It is also suitable for X-ray or neutron diffraction analysis to  
 CC determine three-dimensional structure of mutant PKS and to design  
 CC additional mutants. The crystallisation can serve as further  
 CC purification. Because the synthase may crystallise in more than one  
 CC crystal form, the structural coordinates of alpha-carbon atoms of an  
 CC active site determined from a synthase or its portions are useful to  
 CC solve structure of other crystal forms of synthases. The present sequence  
 CC represents a mutant polyketide synthase of the invention. Note: The  
 CC present sequence is not shown in the specification but was derived from  
 CC the wild-type PKS sequence given on page 13 (see ABG31832)

XX Sequence 389 AA;

Query Match 99.5%; Score 1988; DB 5; Length 389;

Best Local Similarity 99.5%; Pred. No. 3.3e-202;

Matches 387; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MVSVEIRKAQRAEGPATILAIQTANPANCVEQSTYDFYFKITNSEHKTLEKQFQRC 60  
 DB 1 MVSVEIRKAQRAEGPATILAIQTANPANCVEQSTYDFYFKITNSEHKTLEKQFQRC 60

QY 61 DKSMIKRMYLITEILKENPNVCEYMAPSLDARQDMVVEVPRIGKEAAVKAKEWGQP 120  
 DB 61 DKSMIKRMYLITEILKENPNVCEYMAPSLDARQDMVVEVPRIGKEAAVKAKEWGQP 120

QY 121 KSKITHLIVCTTSGVDMFGADYQTLKGLRPVVKRYMYQQGXFAGGTVLRLAKDLAEN 180  
 DB 121 KSKITHLIVCTTSGVDMFGADYQTLKGLRPVVKRYMYQQGXFAGGTVLRLAKDLAEN 180

QY 181 NKGARLVVCSVTAFTFRGSDTHLDSLVGQALFGDGAALIVGSDPVPPEIEKPIFENV 240  
 DB 181 NKGARLVVCSVTAFTFRGSDTHLDSLVGQALFGDGAALIVGSDPVPPEIEKPIFENV 240

QY 241 WTAQTIAPDSEGAIDGHLREAGLTFLHLKDVPGIVSKNITKALVEAFEPGLGSDYNSIFW 300  
 DB 241 WTAQTIAPDSEGAIDGHLREAGLTFLHLKDVPGIVSKNITKALVEAFEPGLGSDYNSIFW 300

QY 301 IAHPCGPAILDQVEOKLALKEPKMNATREVLSEYGNMSSACVLFILDEMRKKSSTONGLKT 360  
 DB 301 IAHPCGPAILDQVEOKLALKEPKMNATREVLSEYGNMSSACVLFILDEMRKKSSTONGLKT 360

QY 361 TGEGLWGVLFPGFGLTIETVLRSAI 389



```

Db      361 TGEGLWGLFGFGPGGLTIETVLRVAI 389
|||||
RESULT 15
ABG31820
ID      ABG31820 standard; protein; 389 AA.
XX      AC      ABG31820;
XX      AC      05-NOV-2002 (first entry)
XX      DE      Polyketide synthase mutant #20.
XX      KW      Polyketide synthase; enzyme; protein coordinate data; PKS; resveratrol;
XX      KW      chalcone; crystal structure; mutant; muten.
XX      OS      Mammalia.
XX      OS      Synthetic.
XX      FH      Key      Location/Qualifiers
XX      FT      Misc-difference 269
XX      FT      /note= "Wild type Lys substituted by Gly"
XX      PN      WO200257418-A2.
XX      PD      25-JUL-2002.
XX      PF      14-DEC-2001; 2001WO-US048523.
XX      PR      15-DEC-2000; 2000US-0255811P.
XX      PA      (SALK ) SALK INST BIOLOGICAL STUDIES.
XX      PI      Noel JP, Austin MB, Bowman ME;
XX      PR      WPI; 2002-590730/63.
XX      PT      Producing mutant polyketide synthase (PKS) by comparing crystal structure
XX      PT      of wild-type (wt) PKS with crystal structure of second PKS and
XX      PT      substituting amino acids of wt PKS with amino acids at homologous
XX      PT      positions in second PKS.
XX      PS      Claim 4; Page; 243pp; English.
XX      CC      The invention relates to a method of producing mutant polyketide synthase
XX      CC      (PKS) comprising: (a) comparing a crystal structure of a wild-type PKS
XX      CC      with a crystal structure of a second PKS; (b) substituting one or more
XX      CC      amino acids of the wild-type PKS with the amino acid residues at
XX      CC      homologous positions in the second PKS; and (c) producing the mutant PKS.
XX      CC      The method is useful for producing a mutant polyketide synthase, and for
XX      CC      altering the activity of PKS, where the altered activity results in the
XX      CC      formation of the product of the second PKS instead of the product of the
XX      CC      wild-type PKS. Preferably, the altered activity results in the formation
XX      CC      of resveratrol instead of chalcone. The crystalline form of PKS is useful
XX      CC      for determining the position of specific alpha-carbon atoms and R-groups
XX      CC      associated with it, comprising the active site, in three-dimensional
XX      CC      space. It is also suitable for X-ray or neutron diffraction analysis to
XX      CC      determine three-dimensional structure of mutant PKS and to design
XX      CC      additional mutants. The crystallisation can serve as further
XX      CC      purification. Because the synthase may crystallise in more than one
XX      CC      crystal form, the structural coordinates of alpha-carbon atoms of an
XX      CC      active site determined from a synthase or its portions are useful to
XX      CC      solve structure of other crystal forms of synthases. The present sequence
XX      CC      represents a mutant polyketide synthase of the invention. Note: The
XX      CC      present sequence is not shown in the specification but was derived from
XX      CC      the wild-type PKS sequence given on page 13 (see ABG31832)
XX      SQ      Sequence 389 AA;

```

```

Query Match      99.5%; Score 1987; DB 5; Length 389;
Best Local Similarity 99.5%; Pred. No. 4.3e-202;
Matches 387; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Search completed: April 22, 2005, 21:41:19  
Job time : 167 secs

*This Page Blank (usptn)*